THE BELL SYSTEM
TECHNICAL JOURNAL

VOLUME XLIV OCTOBER 1965 NUMBER 8

The 2A Line Concentrator

By U. K. STAGG

(Manuscript received May 4, 1965)

The 2A Line Concentrator is a supplement to existing switching systems and provides concentration of line circuits up to a range of 1000 miles. It has been designed to accomplish this range extension while requiring a short (150–200 milliseconds) concentrator work time. An over-all description and a discussion of the initial application are presented.

I. INTRODUCTION

The 2A Line Concentrator is a high-speed, medium range system designed for use in the No. 5 crossbar system to provide range extension with more rapid connection of lines than is now possible. An ac signaling technique and solid-state logic circuits are combined with relay and crossbar switching circuits to provide a concentrator system with a range of about one thousand miles and a call setup time of approximately 175 milliseconds. These are improvements over the existing 1A Line Concentrator[1] (LC1A) system which is range limited to about twenty-five miles and the 1A Line Concentrator system modified for extended range with a call setup time of about 1.8 seconds.

II. OBJECTIVES

2.1 *Application*

The 2A Line Concentrator (LC2A) system was designed principally to provide concentration of lines serving TWX equipment. Although the LC2A development centered about requirements derived from its

proposed application to data systems, its use is not limited to this area. Whenever concentration of lines over a distance greater than that served by the LC1A would become economically attractive, the LC2A may be applied. Several of these systems have been or are being installed at present to provide concentration of dial TWX lines.

2.2 *Development Objectives*

The initial objectives of the LC2A development were essentially as set forth below. As in most systems, some of these objectives changed during the course of the development; however, the statements are the objectives as they have been met.

2.2.1 *Number and Type of Lines Served*

Each LC2A system may serve up to 156 individual line customers. These lines are served through two remote circuits associated with one control circuit over a maximum of 32 concentrator trunks (16 trunks per remote circuit). One hundred and sixty line terminations are provided; however, two line terminations associated with each remote circuit are reserved for maintenance purposes. Each group of 78 lines (maximum) in a remote circuit has full access to any of the 16 trunks (maximum) over which it is served. Facilities are provided for operation with less than the full complement of lines or concentrator trunks, and an "all trunks busy" condition in one remote circuit does not affect the traffic handled by the other remote circuit.

2.2.2 *Customer Loop Length*

The total external loop resistance from the remote circuit to the customer's equipment cannot exceed 2795 ohms. No padding or adjustment on the individual line circuits is required.

2.2.3 *Delay Disconnect and Trunk Preselection*

The LC2A system incorporates delay disconnect, a feature which provides that all but one of the concentrator trunks serving each remote circuit shall remain cut through to the last line served, the remaining trunk becoming the preselected trunk for use on the next call. The choice of the trunk to be disconnected is changed by the trunk preselection circuit on every call, thus preventing a single bad trunk from putting the trunk group serving a remote circuit out of service. This feature is feasible primarily as the result of development of magnetically latching hold magnets for the crossbar switch.

2.2.4 *Line Preference*

To prevent a single line from holding other lines out of service and to spread seizures among all lines when simultaneous requests are made, a line preference circuit which changes preference on each call is used.

2.2.5 *Signaling*

The establishment or disconnection of a path from a line to a trunk, thence into the No. 5 crossbar network, is controlled by signals passed between the remote and control circuits. To provide for the range over which the LC2A must operate, ac signaling is used. The mode of this signaling system is frequency shift pulsing (FSP) also referred to as frequency shift keying.[2] Information is transmitted at a rate of 200 bits per second. Two narrow-band channels are used, one for each direction of signaling over a four-wire transmission path between the remote and control circuits. Both of these bands are within the audio range. Specifically, for signaling from control to remote circuit 2125 ± 100 cps is used, and from remote to control circuit 1170 ± 100 cps. Nominal adjustments of the power levels of these signals are made such that the input to the receiving end is −18 dbm.

2.2.6 *Line and Trunk Switching*

As noted in Section 2.2.1, each control circuit may have associated with it two remote circuits. At each remote circuit is a switching network consisting of four 200-point, six-wire crossbar switches. A similar network is provided at the control circuit for each remote circuit. The operation of these networks is identical at either remote or control circuit. Magnetically latching hold magnets are provided on these crossbar switches to minimize the current drain after a crosspoint is closed, since fifteen of the sixteen available trunks of each remote circuit normally remain connected to the last line served. An advantage of this feature is the ability to maintain a connection of lines to the No. 5 crossbar network in the event of a power interruption at either the remote or control circuit location. During such an interruption, calls cannot be switched; however, this feature precludes the necessity of reestablishing all connections and assures that service transmissions in progress at the time of the interruption will not be preempted by other calls when power is restored. Still another important advantage accrues from this feature — reduction in the average time to connect a customer's station equipment into the No. 5 crossbar

network. Since some of the lines will be left connected to the trunks last used; any calls originating from or destined to terminate to these lines will not have to be switched by the concentrator. In the case of a service request, an off-hook signal will be passed immediately to the No. 5 crossbar office over the concentrator trunk and will cause no action in the concentrator other than that of marking the trunk busy. Similarly, a terminating call will apply ringing and mark the trunk busy with no further concentrator action. The net result of this feature is zero concentrator work time in establishing the required connection.

2.2.7 *Maintenance and Reliability*

Trouble detection and indicating facilities have been provided at both the remote and control units. When trouble occurs at either unit, indications of the nature of the trouble encountered, the line and concentrator trunk identification associated with the trouble, and indications of the call progress are recorded. At the remote circuit, a lamp panel records the call information while at the control circuit the call failure information is punched onto a standard No. 5 crossbar trouble recorder card.

Two different methods of recording call failure information are used since the remote circuits may be located in any type of central office, while the control circuit is always located in a No. 5 crossbar office equipped with a trouble recorder.

Trouble location and verification not only includes the normal techniques associated with electromechanical systems but is extended to facilitate trouble analysis of the solid-state circuits as well. Test points are provided on a group of terminal strips connected to the output of every transistor circuit with the exception of a small number of the ac signaling circuit elements.

Location and verification of a trouble is followed by clearing the cause and replacement of the printed wiring board(s) found to be defective.

Trunks connected to lines found to be in a permanent signal condition are capable of being restored to service from the control circuit end. A special maintenance call — permanent signal service denial — allows disconnection of the trunk from the troubled line and places the line in a cutoff condition, thereby preventing it from reseizing the restored or other trunks. When the permanent signal condition is cleared from the line, it may be restored to service from the control circuit.

Two line appearances are reserved for maintenance purposes. One appearance is required for making operation tests and for maintenance and transmission tests from the remote or control circuit. A second test appearance is necessary to provide loop around transmission testing facilities. Any trunk may be selected for testing, but from the control circuit only.

2.2.8 *DC Power Drain*

The remote and control circuits are combinations of relay and solid-state switching devices. The electromechanical devices, with the exception of a small number of dry reed relays, are operated from central office battery. The solid-state devices and the aforementioned reed relays (operated through conducting transistors) are supplied through a dc-to-dc converter which operates off central office battery to provide a regulated +12-volt source. The total current drain for the remote and control circuits is shown in Table I.

III. SWITCHING PLAN

3.1 *General*

The switching networks at the remote and control circuits are identical configurations of crossbar switches interconnected to provide concentration of customer's lines over concentrator trunks. This crossbar switch network is the heart of the LC2A system, and all functions of the remote and control circuits have the single over-all objective of providing proper control of this network.

3.2 *System Network*

Fig. 1 is a schematic indicating the association of a two remote circuit LC2A system with the network environment in which it is designed to operate. The control circuit is associated with a No. 5 cross-

TABLE I — CURRENT DRAIN

Circuit	State	Central Office Current Drain
Control	Idle	4.0 amps.
Control	Active*	0.9 amps.
Remote	Idle	3.3 amps.
Remote	Active*	5.8 amps.

* Figures obtained by averaging over all calls.

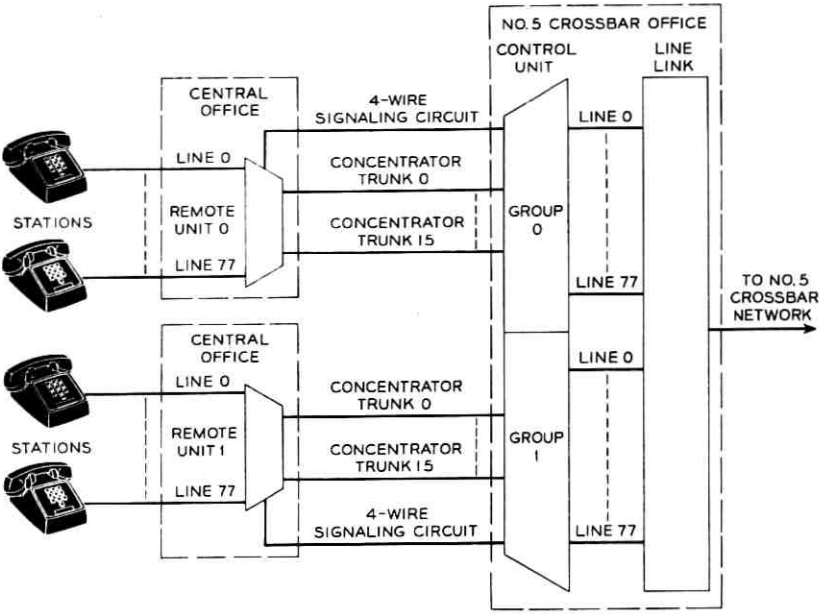


Fig. 1 — System network.

bar central office. The line appearances of the concentrator control circuit are connected to the line relays of the line link frame. Two groups of concentrator trunk circuits interconnect the control circuit with each of two remote circuits, one group of trunks serving each. A four-wire signaling circuit is extended between each remote and control circuit, providing the signaling facilities required for establishing or disconnecting paths through the crossbar switches. The trunks are switched on a two-wire basis through the concentrator crossbar networks; however, the signaling paths are not. The facilities constituting the concentrator trunks and signaling paths differ only in that the transmission trunks (trunks 0-15 of each group) need not be four-wire circuits, while the signaling paths must be to provide the required signaling capabilities. These facilities may be metallic circuits with or without E-type repeaters or carrier.

The remote circuits may be located at a variety of central office types. The choice is a matter of convenience only, since the remote circuits are associated with the central office for the purposes of obtaining central office battery and interconnecting to its alarm system. It is not in any way associated directly with the switching network of the office in which it is located. Customers' line circuits are brought

directly to the remote circuit and associated with the concentrator crossbar switch line appearances. These line circuits, as indicated, may be either metallic pairs, four-wire voice-frequency repeater facilities, or carrier facilities with single-frequency signaling. Though each remote circuit is indicated to be located in different offices, there is no restriction of this nature, and where more than seventy-eight lines are to be concentrated, a second remote circuit may be used at the same location.

3.3 *Concentrator Switching Network*

To reduce the probability of the need for more than one remote circuit at a given location, and to take advantage of the greater efficiencies of larger trunk groups, a larger number of lines and a larger trunk group have been provided than in previous concentrator applications.

3.3.1 *The Network*

The networks at the remote and control circuits are comprised of four 200-point six-wire crossbar switches as shown in Fig. 2. Associated with each vertical of each switch is a customer's line circuit (remote circuit) or an appearance on the line link frame of a No. 5 crossbar office (control circuit) with the exception of verticals 78 and 79. These are the appearances reserved for test and maintenance; they are not associated with line circuits but rather terminate on other equipment within the concentrator system. The sixteen concentrator trunks are multipled to each of the four crossbar switches, corresponding levels of each switch serving the same trunk. In this fashion, all eighty line appearances have full access to all sixteen concentrator trunks. When a connection of line to trunk is being established, the corresponding select magnets of all switches are operated. Following operation of the single hold magnet associated with the line appearance, only one set of crosspoints is operated and all select magnets are released.

3.3.2 *Crossbar Switch Arrangement*

In order to terminate sixteen trunks on a crossbar switch provided with only ten horizontal levels, as indicated in Fig. 2, a scheme allowing the sharing of a level by two trunks has been used. (This is similar to the technique used on the trunk link frame of No. 5 crossbar.) By using levels 0-7 for the trunk appearances and levels 8 and 9 as "steering" levels for selecting one of the two trunks sharing the same level,

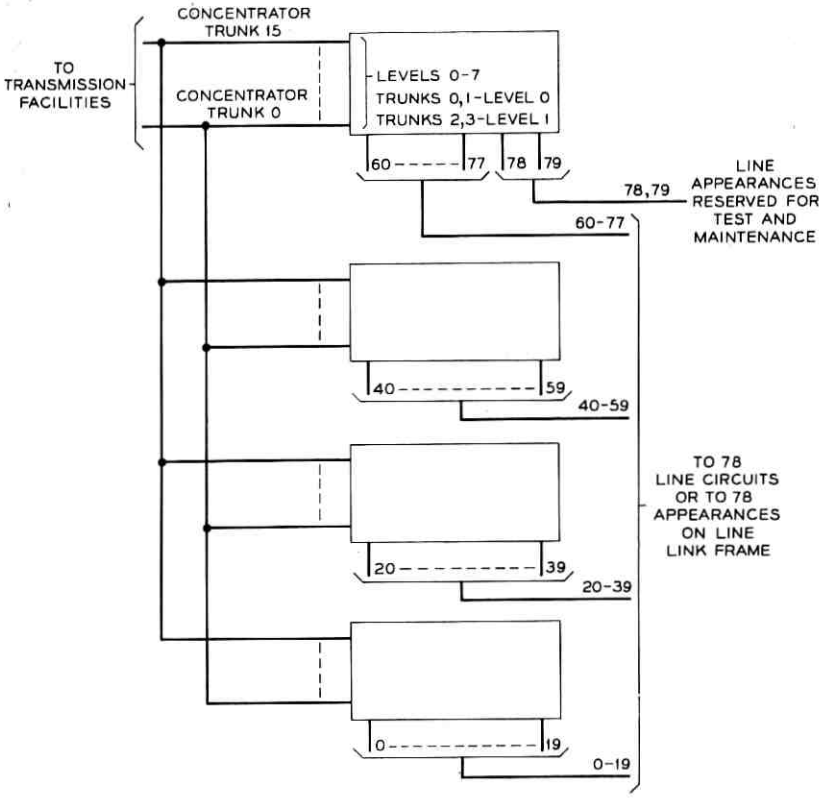


Fig. 2 — Four 200-point, six-wire crossbar switches as employed in LCZA remote or control circuit.

sixteen trunks are made available to each of the twenty line appearances of the switch. Fig. 3 illustrates the manner in which this is accomplished for one vertical unit. Only two trunks are shown, occupying level 0, the remaining trunks appearing in the same fashion on levels 1-7. (Fig. 3 shows a vertical unit for the remote circuit. A minor difference exists between this and the control circuit configuration, but it is in the disposition of the sleeve leads after switching. In general the figure is the same for the control circuit.)

3.3.3 *Trunk Select Steering*

To provide switching of two trunks per horizontal level of the crossbar switch, a minimum of four contacts per crosspoint are required — one for each tip and ring of the two trunks. In addition, a

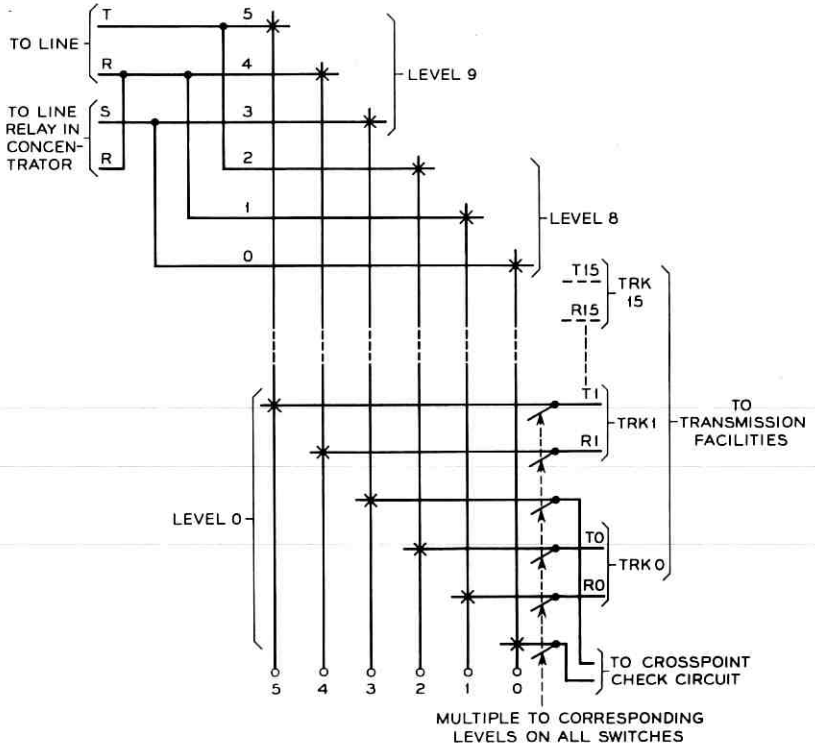


Fig. 3 — Remote unit switch vertical unit.

sleeve lead per trunk is provided for supervision and checking purposes; hence a total of six contacts are required per crosspoint. When connection of a line to a trunk is required, action within the concentrator identifies which line is to be switched to which trunk. Identification of the trunk causes operation of two select magnets on each switch — one corresponding to the level on which it appears and the other for selecting one of the two trunks sharing that same level. Fig. 3 shows two trunks, TK0 and TK1, sharing level 0. TK0 is wired to positions 0, 1, 2 and TK1 is wired to positions 3, 4, 5 of all vertical units of all switches. When the crosspoint at level 0 is operated, all six contacts are closed and both trunks are connected to the vertical unit. The operation of the select magnet for level 8 or 9 "steers" the connection of the line associated with this vertical to the proper choice trunk since only three contacts are provided at these crosspoints. For example, suppose it is desired to connect TK0 to the line. Select magnets 0 and 8 are

operated, followed by the operation of the hold magnet, causing closure of six contacts at level 0 and 3 contacts at level 8. TK1 is connected to the vertical unit also, but no path exists to the line since the contacts at level 9 have not been operated. Note that levels 8 and 9 are not multipled to any other vertical unit of the same or other switches since this is the point of association of the lines with the vertical units.

IV. SIGNALING AND CONTROL

4.1 *Signaling System*

Frequency shift transmitters and receivers are used in both the control and remote circuits. Two separate frequency bands are used: the lower (f_1) to signal from the remote circuits to the control circuit and the upper (f_2) to signal from the control circuit to the remote circuits. In the f_1 band the frequencies corresponding to mark and space signals are 1270 cps and 1070 cps, respectively. Mark and space signals in the f_2 band are 2225 cps and 2025 cps, respectively.

As used in the LC2A, one of the two frequencies is present at all times with one exception. The exception is removal of both frequencies from the transmission facilities to effect the release of the concentrator. This occurs in the normal sequence and is detected by a signal-present detector circuit. If this absence occurs when the concentrator is idle, an alarm is given to indicate that the signaling circuit has failed.

When the concentrator is idle, a continuous spacing signal is transmitted from the remote to the control and the control to the remote circuit. Transmission of a message is initiated by a start pulse of mark frequency, five milliseconds in duration. This message may be started at either the remote or control circuit. Part of the receiver circuit at each end is a guard interval timer (GIT) which has the function of measuring the duration of all mark frequency signals. If the mark frequency signal is of a minimum duration of 4.06 milliseconds, the GIT recognizes the signal as a legitimate mark signal and, in the case of the start pulse, starts the clock at the receiving end. The purpose of this timer is to assure that components of the proper frequency from noise hits as produced by lightning or induced into the transmission path because of other environmental conditions do not cause false starts of the concentrator. This timer circuit is designed to respond only to a continuous signal. A series of short noise bursts with proper components will not build up to give a false start because of the fast recycle time of the timer.

4.2 *Synchronization*

The signaling scheme of the LC2A requires no special synchronization or framing pulses, even though the clocks at the remote and control circuits are independent once set into operation. In regard to the framing, no indication of message length or word length is required, since in all instances the receiving unit "knows" that the transmitted message at any time is of a certain bit length (either 11 or 16 bits), that each word is composed of five bits, and that every message is preceded by a single start pulse. The message length is determined by the point of origination (remote or control circuit) and the signaling stage of the call. All information is transmitted as binary coded two-out-of-five words.

When a start pulse is received, the output of the 4.06-millisecond GIT triggers a 3200-pps clock circuit into operation. This clock drives a four-stage binary counter, which acts as a frequency divider producing a 16-pulse string. Each pulse in this string recurs at a rate equivalent to the driving clock repetition rate (3200 pps) divided by the number of states of the binary counter (16), or at 200 pps. Three of these pulses are used to trigger the shifting, writing (sampling incoming data) and bit-counting circuits. The GIT output also sets this frequency divider to a predetermined state corresponding to the elapsed time as if the clock had started coincidently with the leading edge of the start pulse. The 4.0625-millisecond interval corresponds to 13/16 of 5 milliseconds; hence the counter is set to state 13. In this manner the receiving clock and divider circuits are set into initial synchronism with the incoming signal. Since the clock at the transmitting end is running independently of the receiving circuit clock, and since the repetition rate of these clocks is subject to a $\pm$ 1 per cent tolerance, the two clocks could be as much as 2 per cent out of synchronization. This would indicate that some form of synchronization is in order. Here advantage is again taken of the knowledge of the message structure. Since each word is composed of two mark and three space characters, a transition from space to mark must occur after the initial transition caused by the start pulse — at worst, nine bit positions later. Thereafter, at most a length of ten bit positions is the longest duration between transitions (in the case of a 3-word, 16-bit message).

By positioning the shift, sample and count pulses appropriately in the 16 available positions and by resynchronizing the frequency divider binary counter producing these 16 positions every time the GIT pro-

duces an output, no information is lost even when the two independent clocks are out of synchronization by the maximum 2 per cent allowed.

4.3 *Solid State Logic, Memory and Control*

In addition to applications in the ac signaling circuits and the clock circuits described above, solid-state devices are employed in memory elements, pulse-shaping circuits and logic gates, and as relay drivers. The memory elements are used to: (1.) indicate call progress, (2.) remember the originating circuit identity (control or one of two remote circuits), (3.) indicate to the logic control circuit the originating circuit identity, (4.) temporarily store the message when transmitting or receiving, and (5.) cause operation of relay driver gates to pass the received message to the relay circuits. The pulse-shaping circuits are monostable multivibrators used to stretch or shrink pulses obtained from the clock and frequency-divider circuit. The logic used to control the start, stop, sequencing and preference functions of the solid-state circuits is composed entirely of transistor-resistor logic (TRL) gates. These same gates are connected in re-entrant configurations to construct the binary cells and the monostable and free-running multivibrators. In addition, the same circuit used to perform logic operations is used to operate the dry reed relays in the solid-state-to-electromechanical interface. When used as a relay driver, the relay winding is substituted for the normally provided load resistor of the TRL gate. The reed relays employed were specially coded for this application, operating on 40 milliamps at 12 volts.

Throughout the solid-state part of the LC2A only one code of transistor has been used — the 16A. In the ac signaling circuit it operates as a linear amplifier and as the active element in oscillator circuits. Its application in the remainder of the solid-state circuits is in the role of a switch.

V. SYSTEM DESCRIPTION

5.1 *General*

This section described the salient features of the LC2A system and includes a discussion of the signaling method and solid state — electromechanical interface. A summary of the system characteristics is given in Section 5.5.

5.2 *Line Selection*

Simultaneous originating or terminating requests must be queued and served preferentially. The concentrator may be simultaneously summoned by more than one line at the remote circuit to connect these lines into the No. 5 crossbar network. Similarly, simultaneous termination requests from the No. 5 network to more than one line (two or more different calls) may be placed on the concentrator. Still another condition is simultaneous requests for serving a line from the remote circuit (service request call) and a request for terminating to a line (terminating call) from the control circuit. Since the LC2A can serve only one call at a time, a preference and lockout arrangement must be incorporated in the design. Where the simultaneous bids are originated at the same point (remote circuit or control circuit), an electromechanical technique suffices. What must be accomplished is recognition of the lines requesting service, determining which is to be served, remembering the identification of the served line and preventing others, waiting to be served, from interferring.

These ends have been accomplished in the LC2A by combining the classical lockout chain circuit with the circuit used to translate the operation of a line relay into a corresponding decimal number. The operate paths of ten relays used to determine the units digit of this number are chained through back contacts of these same relays. In similar fashion, the operate paths of the relays used to determine the tens digit are also chained. By providing a W-Z relay circuit, the entry point into, and preference through, each of these chains is reversed on alternate calls. Through this mechanism three things are accomplished: (1.) simultaneous requests result in one-only identification, (2.) line preference distribution tends to be smoothed by alternately preferring high-numbered and low-numbered lines, (3.) a single line in a trouble condition cannot block permanently other lines from service. Fig. 4 shows the line lockout and preference circuit for the units digit of the line number at the remote circuit.

In the event of simultaneous seizures of the remote circuit (for a service request) and the control circuit (for a terminating call) both units will initiate signaling to the other. This results from the nature of the signaling scheme (discussed in Section 4.4). By a suitable arrangement of solid-state logic and control, the LC2A system forces itself to abandon the call originating at the remote circuit and continues processing the control circuit originated terminating call. This choice was

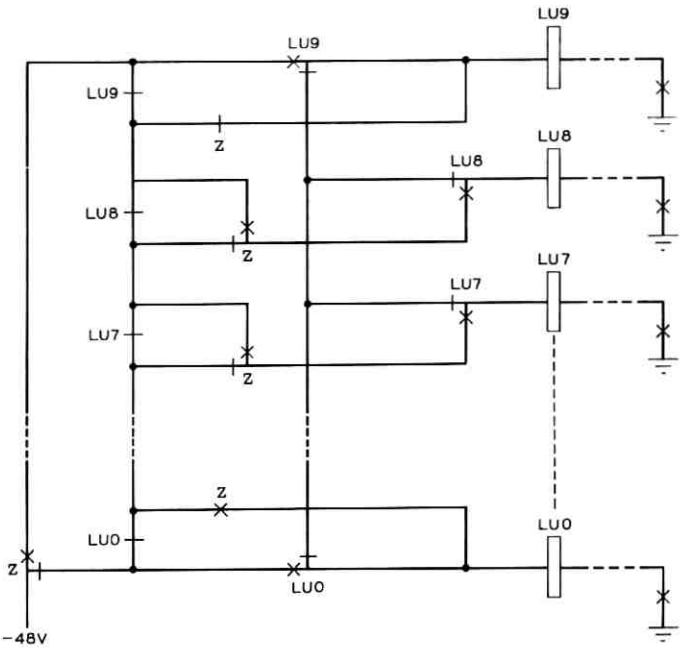


Fig. 4 — Line lockout and preference circuit.

made in order that a call which has already been through several stages of switching in the No. 5 crossbar network will not be denied completion or forced to wait for completion of a service request call which has not been switched at all.

Another possibility which must be guarded against is simultaneous seizure of the control circuit by both remote circuits. Since the control circuit is seized at the time it recognizes that a remote circuit is sending a message, the decision of which remote to serve must be rapid in order that information from the unserved circuit will not interfere with the valid information from the served circuit. In addition, the unserved remote must be so informed to prevent it from giving a false alarm. Fig. 5 shows the preference arrangement which, in the event of information arriving simultaneously at the control from both remotes, causes both flip-flops — indicating which remote circuit is to be served — to be set. However, an arrangement in the logic forces one (arbitrarily selected) flip-flop to be reset and disables the receiving circuit associated with that remote circuit. Under normal conditions the set state either flip-flop disables the receive circuit for the other remote circuit.

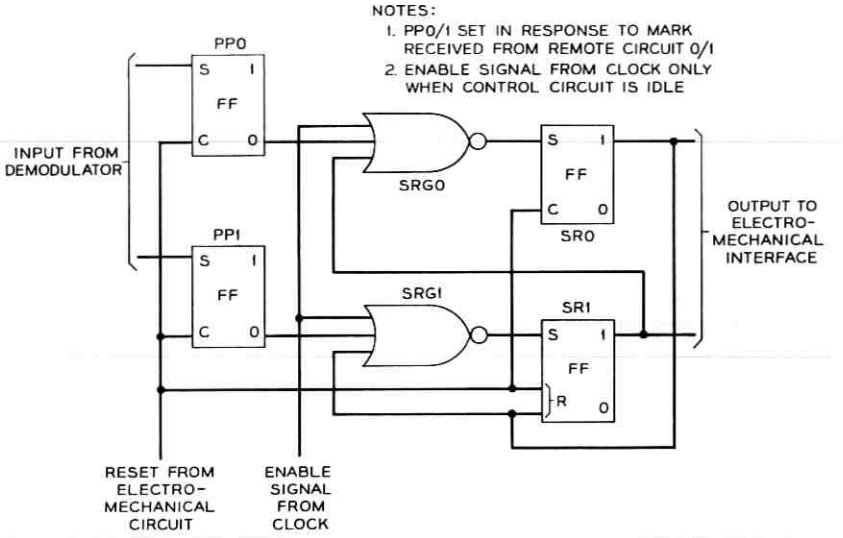


Fig. 5 — Remote circuit preference and lockout circuit at control circuit.

5.3 *Trunk Selection*

The trunk to be used on any given call is chosen immediately following the establishment of a previous call if an idle trunk is available, or if not, as soon as one becomes available. Under normal conditions this reduces the average cut through time from customer's equipment to the No. 5 crossbar office at the time connection is requested, since the identity of the trunk to be used is known or "preselected". The identity of the trunk does not have to be passed between the remote and control circuits as part of the information necessary at the time the connection is desired. In fact, before the customer's line has been identified at both ends, the select magnets associated with the preselected trunk are energized at both the remote and control circuits. Identification of the customer's line results in operation of the appropriate hold magnet, thereby closing the proper crosspoints. The identification of the trunk used on any connection is passed between the remote circuit and control circuit after the line has been identified to verify that the same trunk has been selected at both units.

As described in Section 2.2.7, a feature of this concentrator is the ability to leave all but one concentrator trunk connected to the last line served. Naturally, if all trunks are occupied with an active call, no trunks will be disconnected since all will be marked busy. However,

the first circuit to become idle removes its busy indication by releasing a trunk busy (TB-) relay and is immediately selected by the concentrator control circuit as the trunk to be used on the next service call. The selection of this trunk requires storing its identification in the trunk memory of the control circuit, transmitting this identification to the remote circuit, and storing the trunk identity in the trunk memory of the remote circuit. In the normal situation, all trunks will not be occupied, and following the establishment of a call using the preselected trunk, another trunk must be selected out of the pool of idle trunks remaining connected to the last lines served. A trunk preference circuit determines which trunk is to be disconnected *if it is idle*. If the preferred trunk is not idle, the next higher-numbered trunk which is idle is disconnected and its identity stored in the trunk memory circuits. The preference circuit is advanced as the result of each trunk selection operation of the concentrator; consequently each trunk of a group serving a remote circuit is the preferred trunk for preselection once every sixteen calls involving that remote circuit. In this fashion, a trunk which is malfunctioning will not cause blocking of normal operation by preventing selection of other trunk circuits.

In addition to selection of trunks through the preference circuit, a feature has been included to allow manual selection of an idle trunk. This feature is used for maintenance calls and is accomplished by operation of a switch and key combination at the control circuit. Initiation of manual trunk selection overrides the normal trunk selection by releasing the trunk memory of the control circuit and causing transmission of the new trunk identity for storage at the remote circuit. A special code accompanies transmission of the trunk identity, causing the remote circuit to release its trunk memory before storing the new trunk identity.

5.4 *Signaling and Logic Applications*

The FSP (Frequency Shift Pulsing) signaling and the solid-state logic hardware have been provided for the primary purpose of passing the line and trunk identification necessary for establishing a transmission path between a customer associated with a remote circuit and the No. 5 crossbar network at speeds and over distances greater than attainable with the LC1A. A brief description of the different calls handled by the LC2A is presented in the following paragraphs.

5.4.1 *Service Request Call*

A call originating from a customer's equipment associated with the remote circuit is designated a service request call. A request for service

is initiated by an off-hook condition, recognized at the remote circuit by operation of a line relay associated with the customer's equipment. Operation of this line relay results in translation first to a two-digit decimal number, which is stored in a temporary relay memory. The line identification is stored so that in the event of encountering trouble after operation of the crossbar switches at either or both ends but before release of the concentrator, the connection may be released and the concentrator restored to the condition existing before the call started.

After the line identification is stored, a second translation, from decimal to binary two-out-of-five, is performed and the results passed to the shift register circuit. When the information is properly stored, a start signal is passed to the control circuit which enables the clock. As the clock pulses are generated, the information stored in the shift register is passed, one bit at a time, to the FSP modulator circuit. Each output bit of the shift register is an input signal to the FSP circuit for a period of five milliseconds; hence a pulse of five milliseconds duration of the appropriate mark or space frequency is transmitted. Preceding the ten-bit message (two two-out-of-five coded words) is a one-bit start signal; hence, the total message length is eleven bits, requiring fifty-five milliseconds for transmission.

When the clock circuit has generated eleven pulse trains, causing the eleven-bit message to be completely shifted through the shift register to the FSP modulator, the solid-state control circuit disables the clock and starts timing for a response from the control circuit. During the fifty-five millisecond transmission interval, the relay control circuit has caused operation of the crossbar switch network at the remote circuit to connect the identified line to the preselected trunk.

At the control circuit, action is initiated by reception of the start pulse. This pulse enables the clock which generates a string of clock pulses identical to those generated by the remote circuit clock but in synchronism with the arriving pulses. (The clocks at the remote and control circuits are not themselves synchronized because of transmission path delay.) The first mark (the start pulse) is not stored, but subsequent marks are. As the FSP demodulator responds to the incoming signal, it produces output signals of battery or ground accordingly as mark or space frequencies are detected. This output is sampled periodically (every five milliseconds), and if a mark is being received at the time of sampling, the first cell of the shift register is set to indicate the fact. The shift register is then advanced and prepared for the next sampling operation. If a space is received and sampled, no action is taken before shifting of the register.

The receiving clock continues running until an interval of fifty-five milliseconds (corresponding to the time required for reception of the eleven-bit message) has elapsed. The solid-state control circuits then disable the clock and cause gating of all the cells of the shift register into the relay control circuits. The message is immediately checked for two-out-of-five validity, whereupon it starts action for closing the crossbar crosspoints associated with the received line identification and the preselected trunk. Simultaneously with this action, the identification of the line requesting service causes a bridge to be placed on the corresponding circuit to the No. 5 crossbar line link frame, resulting in operation of the line relay. This initiates a dial tone request before the crossbar switch has operated, thus reducing the dial tone delay by an amount equal to the operate time of the crossbar switch hold magnet. Subsequent to operation of the crossbar switch the bridge is removed, with status of the path to the No. 5 crossbar office supervised by the customer's equipment over the concentrator trunk.

As soon as the hold magnet is operated at the control circuit, the number of the trunk used in the connection at the control circuit is translated to a two-word, two-out-of-five message, passed to the shift register and transmitted to the remote circuit in a fashion similar to transmission of the line identity from the remote circuit to the control circuit in the first stage of the call. The remote circuit receives and stores this trunk number just as the control circuit received the line number. The message is checked for two-out-of-five validity, then checked against the stored number of the trunk connected at the remote circuit. If a match results, a verification signal is transmitted to the control circuit to indicate that the information was received and that it corresponded to the trunk number used by the remote circuit. Note that the transmission of the trunk number at this time is not required for establishing the connection but only to verify that both ends used the same trunk. Following this verification, both the remote circuit and control circuit release.

If at any time either unit fails to receive a valid two-out-of-five coded message, its action is to withhold its next transmission. The last unit to transmit starts timing at completion of the transmission, and if a response is not received within the proper interval, preceding action is negated by release of the concentrator and restoration of conditions existing prior to the start of the call.

A second trial feature has been incorporated in the LC2A which allows a second attempt after a timeout. No trouble indications are given unless the second trial fails. In this event a trouble record is made, indicating the nature of the failure, the numbers of the line and

trunk involved, the type of call being processed, and the signaling stage of the call.

In order to inform the unserved remote circuit that it may not serve calls during the time the control circuit is occupied by the first remote circuit, the control circuit removes all signals from the transmission path to the unserved remote circuit. This remote circuit responds by removing its signals to the control circuit and disables itself. This condition obtains until the control circuit reapplies signal to the remote circuit, indicating that it is idle and is available for service. The remote circuit responds to this signal by reapplying its signals to the control circuit, at which time the system is ready to serve another call.

5.4.2 *Terminating Call*

A call destined to terminate at a customer's equipment through the LC2A from the No. 5 network is classified a terminating call. For the LC2A this call originates at the No. 5 line link frame which supplies ground over the sleeve to operate a relay in the control circuit identifying the line. This operation is followed by translation first to a decimal number then to a two-out-of-five coded equivalent. Operation of the LC2A is similar to that described above for a service request call with the signaling roles reversed. The line number is passed from the control to the remote circuit and trunk verification is transmitted from the remote to the control circuit. One major difference in signaling exists, however. Since the remote circuit can only initiate service requests, no information regarding the type of call is passed on a service request call. The control circuit, however, has the ability to originate more than one type of call; therefore, it must signal the remote circuit to inform it of the call disposition. This signaling is accomplished by transmission of a third two-out-of-five coded word representing the type of call. This information is sent with and precedes the remaining ten bits of information. On all calls originating from the control circuit sixteen, rather then eleven, bits are transmitted from the control circuit to the remote circuit. The solid-state control circuits at each unit recognize this and permit the associated clocks to run for the longer duration (eighty milliseconds) rather than cutting them off after the shorter interval used on a service request call.

5.4.3 *Other Control Circuit Originated Calls*

In addition to the terminating call, the control circuit originates the following: (1.) disconnect call, (2.) service denial call, and (3.) test

calls. Each of these is signaled to the remote circuit in the same manner as a terminating call, but the type of call indication indicates a different significance for the ten-bit message which follows the type of call information. Discussion of the operations of these calls is presented rather than the signaling involved.

5.4.3.1 *Disconnect Call.* A feature of the LC2A is the practice of leaving all but one trunk connected to the last line served. The one trunk not connected serves as the trunk for the next call served. When a call is served (service request or terminating), the preselected trunk is used and another trunk, if one is available, must be disconnected from a line and placed in the idle, preselected state. The selection is accomplished by the control circuit which ascertains the identity of the trunk through the trunk preference and select relay circuit. Upon identification of this trunk, the control circuit signals the remote circuit that a disconnect call is being processed and passes the trunk identity to the remote. After the identity is checked for two-out-of-five validity and stored, a relay circuit causes operation of the line relay at the remote circuit of the line connected to the trunk. This is accomplished by grounding the sleeve of the trunk. Operation of the line relay identifies the hold magnet holding the connection, and a current pulse in a direction to overcome the magnetic field latching the vertical unit is applied. The crosspoints release, and a check is made to verify the release through back contacts of the hold magnet. Following the release, the remote circuit transmits the number of the released trunk to the control circuit, where it is verified by checking against the trunk number stored in the control circuit trunk register.

5.4.3.2 *Service Denial Call.* In order to prevent a line at the remote circuit from requesting service, a service denial feature has been incorporated. This feature is necessary to prevent lines which are in a premanent signal or other trouble condition from removing a trunk from service. Service denial is initiated at the control circuit by connecting ground to the sleeve lead at the control circuit through a key. Operation of the key initiates a service denial call and identifies the line. The type of call and line number are transmitted to the remote circuit which in turn energizes the hold magnet associated with the denial line. The hold magnet latches operated and through its back contacts opens the operate path of the line relay. No select magnets are energized; hence, no crosspoints are operated. A check is made that the hold magnet has operated and latched, and verification is returned to the control circuit. In similar fashion, the hold magnet at the control

circuit is also latched, closing no crosspoints and opening the sleeve lead. Hence, terminating calls to the denied service line cannot be served.

A second key at the control circuit allows restoral of the line to service after the trouble condition has been cleared. This key causes transmission to the remote circuit of a release service denial signal and the line identification. The hold magnets are released and checked, and upon release of the concentrator the line is again able to receive or originate calls.

5.4.3.3 *Test Calls.* The LC2A has provisions for making service test calls to test the ability of the concentrator to serve regular service request or terminating calls. In addition, two lines have been reserved for making transmission test calls. Tone supplies (1000 cps, 1 milliwatt) are connected through the test line appearances and, by selection of the type of call, either transmission in one direction or loop around transmission testing may be accomplished. For loop around testing, line appearances 78 and 79 are each connected to a trunk and connected together at the remote circuit. By transmitting from one of these appearances at the control circuit and monitoring the other appearance (at the control circuit) a measurement of the trunk losses may be obtained.

5.5 *Summary of System Characteristics*

Table II summarizes the characteristics of the Line Concentrator No. 2A.

VI. EQUIPMENT FEATURES

6.1 *General*

The LC2A system incorporates some interesting equipment and apparatus applications. These result from the proposed use of the system and the combination of electromechanical and solid-state switching circuits within the same bay of equipment. Some of the choices of equipment arrangement and apparatus selection were results of laboratory testing, particularly to overcome the troubles resulting from noise generated by the electromechanical switches. Both the remote circuit and the control circuit use the same apparatus; and, in general, the equipment arrangements are similar.

TABLE II — SUMMARY OF LC2A CHARACTERISTICS

Size	
Remote:	80 lines (2 reserved for test) with full access to 16 trunks.
Control:	2 remotes (optional) 160 lines 32 trunks
Signaling	
Mode:	4-wire FSP 2/5 coded message.
Frequency:	Remote to control — 1070–1270 cps. Control to remote — 2025–2225 cps.
Range:	Transmission limited to 1000 miles.
Rate:	200 bits per second.
Concentrator Trunks	Metallic or carrier
Switching Network	
Remote:	Four 6-wire, 200-point, magnetic latching crossbar switches. Lines on verticles. HON (Hold Off Normal) contact of crossbar switch as line cutoff. 16 trunks on horizontal/level steering.
Control:	Four switches as above per remote.
Equipment and Apparatus	
Remote:	Central office mounted bulb angle frame.
Control:	Sheet metal frame.
Common to Both:	AK, general purpose, mercury, reed, 286 and B-type relays, transistors, diodes, resistors, capacitors, inductors, transformers, AMPLAS and printed wiring boards

6.2 *Frames and Equipment Arrangement*

The LC2A control circuit is designed to be used only in a No. 5 crossbar office; and, therefore, the frame is of sheet metal construction. Fig. 6 is a photograph of the laboratory model of the control circuit. The right-hand bay contains the two crossbar switch networks required to serve two remote circuits. Located between the networks are two shelves for mounting the printed wiring boards, which contain all but a few of the solid-state circuits. Also mounted in the top shelf, extreme left, is the dc-to-dc converter, which supplies the +12 volt needs of the solid-state circuits. Immediately below the second shelf are two mounting plates containing the test terminal strips for the solid-state circuits and AMPLAS component assemblies mounting diodes, resistors and capacitors associated with the printed wiring boards. The lower plate mounts the transformers and filters for the two four-wire circuits used to signal to the two remote circuits.

The left-hand bay includes the line lockout, trunk preference and

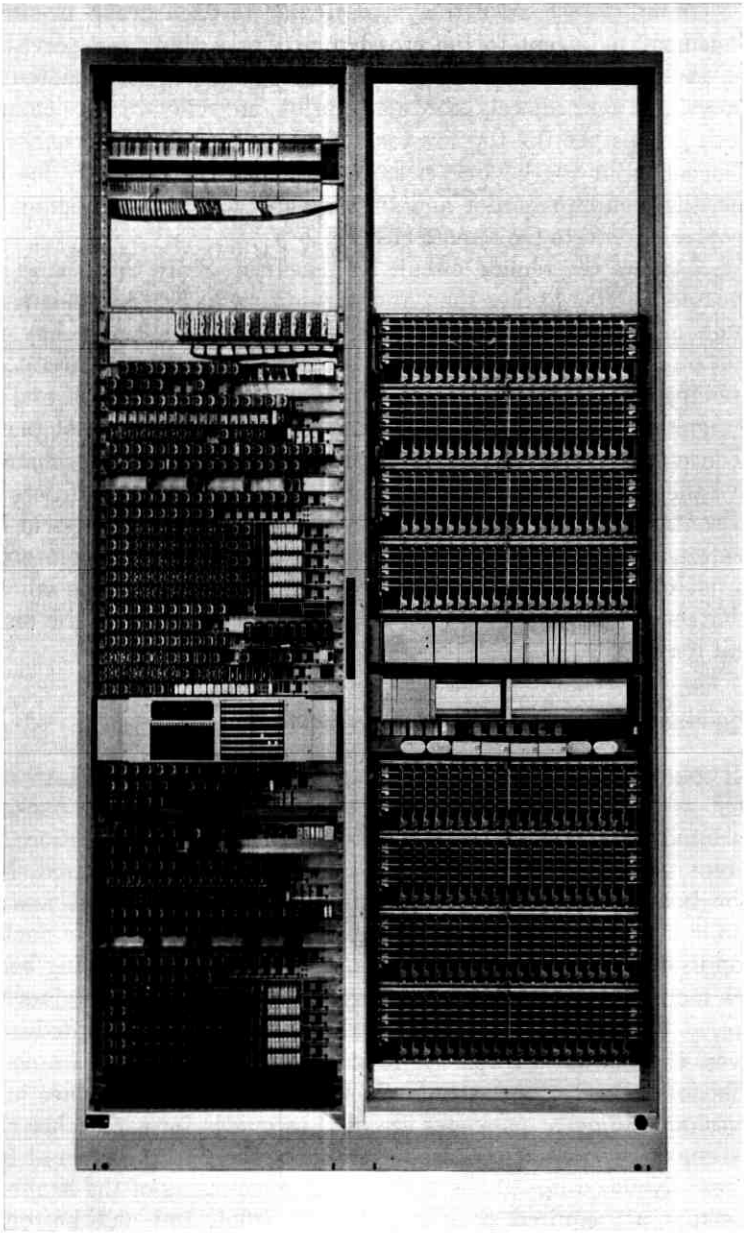


Fig. 6 — Laboratory model of control circuit.

select circuits and control relays particular to each group in similar arrangements adjacent to the crossbar networks that they serve. Between these two groups of relays is a panel provided for maintenance purposes, and immediately above and below are relay circuits common to both groups. At the top are located the terminal strips connecting the concentrator to the line circuits of the No. 5 crossbar line link frame, the trouble recorder and the transmission facilities connecting the control circuit to the remote circuit.

Fig. 7 shows the remote circuit arrangement. Since this frame may be located in offices other than and in addition to No. 5 crossbar, it is of bulb angle construction. All equipment fits on a single bay unit and is similar in description to that located at the control circuit. (In the photograph, lower right-hand corner, are shown two E1L and two E1S signaling units used in laboratory testing. These are not part of the concentrator but were mounted there as a matter of convenience.) Above the mounting plate containing the transformers and filters is a strip containing lamps, jacks and keys. The purpose of this strip is to provide test facilities and a trouble lamp display. The remote circuit does not connect to any trouble recording device within the office. A trouble record is made on a lamp display panel mounted on the remote circuit itself.

6.3 *Solid-State Circuit Packaging and Mounting*

All solid-state circuits with the exception of a few AMPLAS component assemblies mounting a special long-period timer are packaged on printed wiring boards of phenolic or fiberglass construction. The different substrates result from the different strengths required. Most of the boards mount very small, light-weight components, and the phenolic material is satisfactory. Fig. 8 shows a typical logic package at top using the phenolic substrate and one of the ac signaling boards which mounts three relatively massive inductors on the fiberglass substrate, at the bottom. For rigidity and separation when mounted in the shelves, the boards are riveted to an extruded aluminum frame. An Amphenol connector and spring clip hold the packages in place in the shelves, and adjacent packages are used as guides for each other when replacement is necessary. The shelves are of sheet metal, open at front and rear. Notched lips at the rear facilitate mounting of the Amphenol connectors at required positions. At the front, top and bottom, a bronze spring riveted to the frame holds the packages in place in the connectors. In production models, dummy frames have been provided

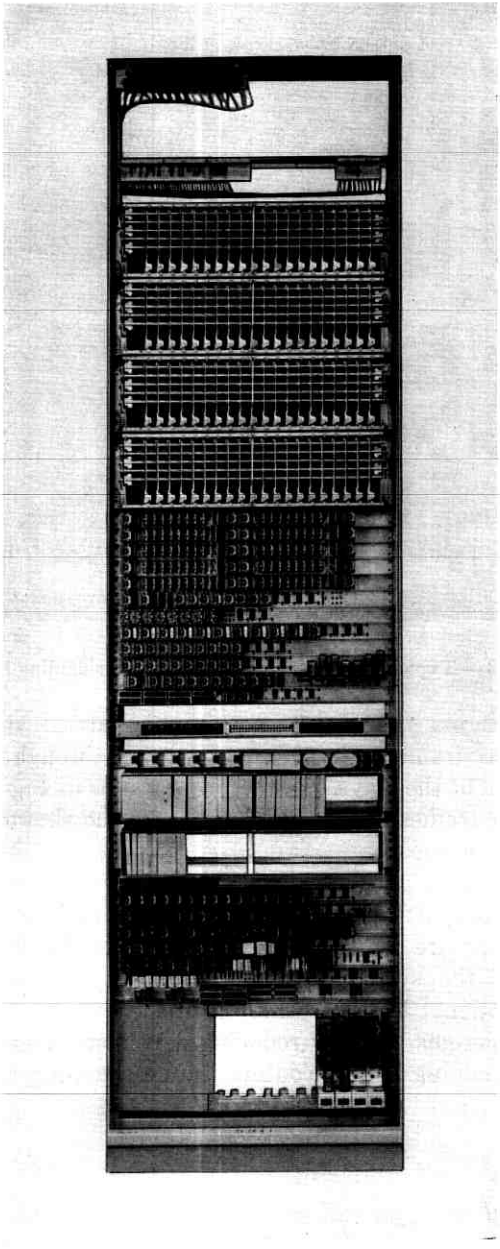


Fig. 7 — Remote circuit arrangement.

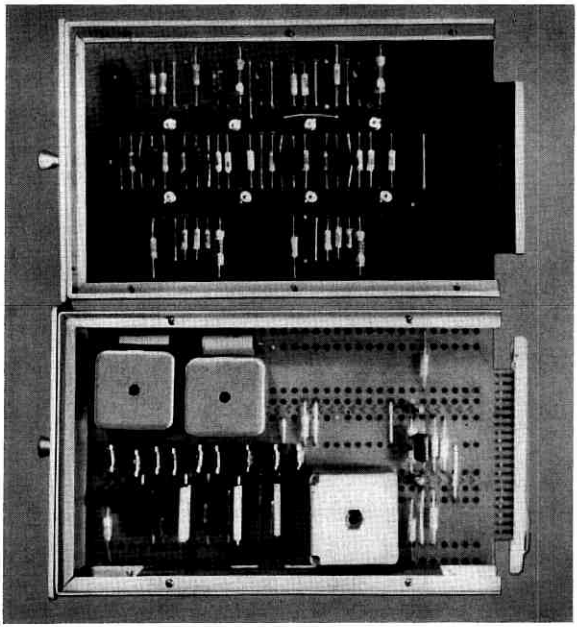


Fig. 8 — Typical logic package (top) and ac signaling board.

at intervals across the shelves to provide lateral stability of the packages. These frame positions are adjustable through slots cut in the top and bottom of the trays. (In the photographs of Figs. 6 and 7 these dummy spacer frames are not present; hence the skewed and variable separation appearance.)

Thirteen different packages have been designed. Seven of these are boards which provide logic gates and bistable and monostable multivibrators. These are interconnected to provide the functional solid-state logic of the system. This method was followed rather than design of functional boards to minimize the number of different circuit package designs, thereby reducing development and maintenance costs. The remaining boards contain the ac signaling circuits and the 3200-pps clock.

6.4 *Noise Interference*

The transient voltages caused by the switching of electromechanical devices create a severe problem in any electronic switching application. Since the packages used in such an application must be interconnected

by surface wiring and local cabling, many points of entry exist for noise voltages. Resistor-capacitor filters applied to the critical points have satisfactorily minimized this problem. A second source of noise interference was determined to be through the ground wiring of the system. Potentials as high as several volts existed in the same run during switching operations, and the wiring was found to support high induced voltages at sufficient distances from frame ground. To overcome this problem a solid copper bar 1 inch by $\frac{1}{8}$ inch has been mounted between the two solid-state shelves, and each package is provided with its own connection to this bar. The bar is maintained at frame ground by adequate wiring to the ground supply at the point where the frame is supplied.

6.5 *Relay–Solid-State Interface*

At some point between the reception of information by the ac signaling circuit and the application of this information to operation of the crossbar network, an interface from transistor to relay logic must exist. In the LC2A this interface problem has been satisfied by using a dry reed relay package specially coded for operation with low-current transistors. The relay operates at 40 milliamperes delivered from a twelve-volt supply. The 16A transistor used in the LC2A has a maximum dc current rating of 50 milliamperes; hence is adequate with margin to switch the reed relay.

As information must be passed from the solid-state to the electromechanical circuit, the opposite direction of information transfer also must be achieved. General purpose wire spring relays, reed relays, etc., have the characteristic of contact bounce associated with dry contacts. This bounce can result in interpretation by the solid-state logic as two or more equally valid sequential signals when only one is meant. To prevent such interpretation, mercury contact relays are used wherever such misinterpretation might result in false operation.

VII. MAINTENANCE AND TEST FEATURES

7.1 *General*

Maintenance of the Line Concentrator No. 2A involves problems normally not encountered in local central offices. The system itself may cross operating company boundaries since operation over a range of up to 1000 miles is possible. Remote circuits may be located in unattended offices. These two factors indicate that coordination of

maintenance personnel at short notice might be difficult if not impossible.

The combination of relay and transistor switching logic places an extra burden on the maintenance forces if straightforward, relatively simple testing facilities and maintenance practices are not provided. Detection and replacement of a single inoperative component in an electronic circuit can be vexing and time-consuming at best, and in some cases may defy the most experienced engineer equipped with sophisticated test equipment for intolerable periods.

To circumvent this need for detailed diagnosis and repair, the maintenance of the LC2A has been simplified as much as practicable. Facilities for establishing test calls from either the remote or control circuit are provided. Trouble recording is included to narrow the range of speculation when trouble shooting. Test points are extended to test terminals from every logic element of the solid-state circuit in order that suspect units may be monitored. Monitoring is done by a simple logic test set which responds to ac signals, dc levels and short duration (microsecond) dc pulses. When a trouble has been localized to a circuit package, the malfunctioning circuit is replaced by simply removing the bad package and plugging in a spare of the proper type. If trouble is localized to the relay circuits, established procedures for relay circuit maintenance are followed.

7.2 *Test Sets*

For routine maintenance and trouble-detection, a test set designed for use on the B1 data carrier terminal is used. The unit uses solid-state circuits arranged to provide low load on the monitored circuit by presenting a high impedance level. When in use as a pulse or level detector, the test set detects on a go, no-go basis the presence or absence of a minimum voltage and displays the result by causing appropriate lamp indications. The set also may be used to obtain a good approximation of the frequency of sinusoidal signals or the repetition rate of nonsinusoidal periodic waveforms. Thus, every part of the solid-state circuit may be monitored and an indication obtained of proper or improper operation.

One of the "gray" areas in the subject of routine maintenance results from the close tolerance which must be held by the clock circuit. As mentioned before, a variation of more than ± 1 per cent in the clock repetition rate may lead to lost calls or a totally inoperative condition. This indicates periodic checks of the clock repetition rate to

insure that tolerance limits are not exceeded. The design of the clock circuit is such that temperature excursions and voltage variations normally occurring in a central office environment will not push operation beyond the limits. What remains an unknown factor is aging of the time-determining components. At any rate, periodic checks of the clock circuit repetition rate with an accurate instrument are required — the required frequency of checks has not as yet been ascertained.

Use of an oscilloscope by experienced personnel may be required under certain conditions where the determination of time relationship of events is necessary to ascertain the trouble. In relay circuits, blocking of key relays and observation of others corresponds to some degree to this use of the oscilloscope. No simple technique is available as a replacement for this method of trouble detection.

VIII. LABORATORY TESTING

8.1 *General*

Laboratory testing of the LC2A was accomplished in three distinct phases. A test to verify the signaling capabilities in the presence of noise was conducted on noise simulator facilities. Following completion of the noise tests, complete system testing of two remote circuits connected back-to-back with the control circuit was conducted. Finally, the system was tested over facilities that were composites of the types of facilities the concentrator would encounter in normal service.

8.2 *Noise Testing*

Testing of the solid-state signaling and logic circuits was conducted to determine the levels of white and in-band impulse noise which would interfere with signaling between two concentrator units. Every valid message combination possible between a remote and control circuit was established; the signaling level was varied beyond the design limits; and noise of varying amplitude and structure was introduced into the signaling channel. The results of these tests showed the signaling capability to be in excess of the requirements, which are:

The operation of the limiter-demodulator, in a noise environment consisting of in-band noise pulses in the range of 48–53 dbRN and at a rate of 35 counts in 30 minutes, is such that the error rate of received information shall be less than one error in 10^5 bits. The range of operation of the limiter-demodulator is -10 dbm to -30 dbm.

8.3 *Laboratory System Tests*

To insure that all features of the LC2A system functioned properly, back-to-back testing was performed. Both remote circuits were wired directly to the control circuit over zero loop trunks, and signaling circuits and load tests were applied. Upon completion of these tests, repeatered facilities were tested by connecting the concentrator trunks through E-type repeaters and reapplying the load tests. These tests were followed by a much more realistic situation. Long-haul facilities were leased from an operating company with both ends of each of two loops terminated appropriately on both remote circuits, and the control circuit and the load tests were once again applied. Testing over these facilities (shown in Fig. 9) was conducted over a period of one month.

As a result of the entire testing program, several minor deficiencies

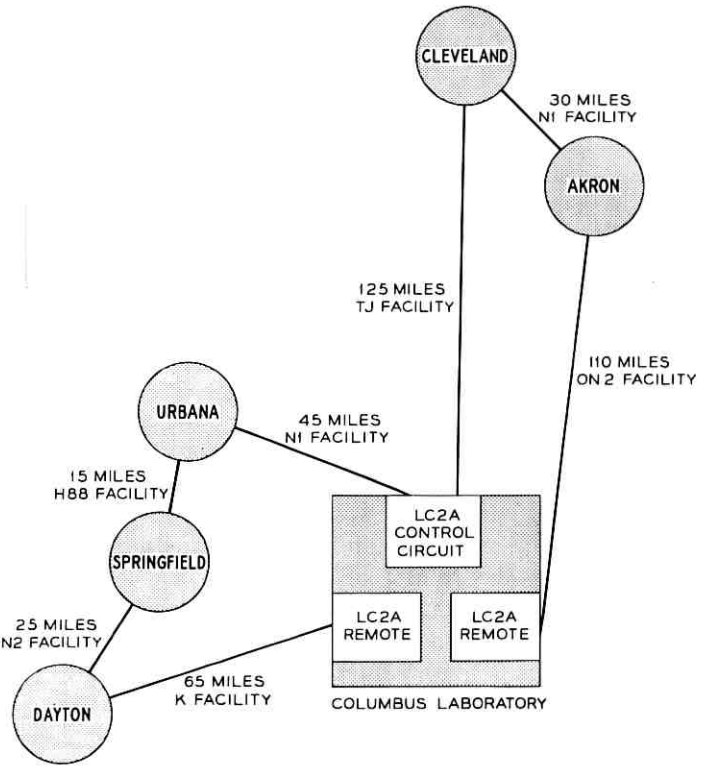


Fig. 9 — Facilities involved in long-haul testing.

which otherwise would not have been detected until after installation in the field were exposed and remedied. Though these shortcomings were minor, the value of the complete testing program undertaken is reflected in the savings obtained through changes made in the shop rather than in the field.

IX. FIELD TESTING

9.1 *General*

Under normal circumstances, a field trial is conducted when a new service or system is introduced to ascertain the adequacy of the design in regard to its performance and customer satisfaction. However, early in the development of the LC2A it was determined that in order to meet the short schedules anticipated, a field trial would not be conducted, but rather an initial installation trial conducted cooperatively by Bell Telephone Laboratories and the operating company would be substituted. Though the time requirement was eased at the completion of the development, it was decided to follow through with the initial installation trial plan.

The first LC2A was installed in Wisconsin and placed in service on April 30, 1964. The control circuit is located in Milwaukee with its two associated remote circuits located in Appleton and Green Bay, Wisconsin. The distances (air miles) between the control and remote circuits are approximately 100 and 125 miles, respectively. The facilities over which the concentrator trunks and the signaling circuits operate are ON carrier. All three circuits (both remote and the control) are located in toll equipment rooms at their respective locations.

Initial system testing was begun in January, 1964, and completed in February, 1964. Load testing similar to that accomplished in the laboratory was first applied, followed by placing dial TWX customers in service on the system after satisfaction that proper operation was being achieved.

The initial installation trial terminated in March, 1965. During the eleven month trial period, status reports, line and trunk development and trouble reports were made by the operating company.

X. CONCLUSION

The LC2A system has met the objectives of providing a longer range for concentration of customers' lines without a penalty of excessive concentrator work time. This has been done through the com-

bination of solid-state logic circuitry, electromechanical switching networks, and frequency shift signaling technique. It was not intended nor should it be construed to be the ultimate in concentrator systems. Rather it can be viewed as the specialized forerunner of a more sophisticated system which, by taking advantage of past, current and future developments in switching techniques, may provide a much wider range of general application.

REFERENCES

1. Krom, M. E., The 1A Line Concentrator, Bell Laboratories Record, *40*, September, 1962, pp. 297–302.
2. Bennett, W. R., and Rice, S. O., Spectral Density and Autocorrelation Functions Associated with Binary Frequency-Shift Keying, B.S.T.J., *42*, September, 1963, pp. 2355–2385.

Wideband Data on T1 Carrier

By L. F. TRAVIS and R. E. YAEGER

(Manuscript received May 18, 1965)

The T1 carrier repeatered line is the Bell System's first high-speed digital transmission facility. Although developed primarily for the transmission of analog information in the form of processed voice frequency signals, its potential use as a short-haul data facility has stimulated considerable interest and study. T1 carrier data terminal designs are described for various general types of applications, and consideration is given to certain problems concerning the transmission of data-type signals over the repeatered lines. The applicability of the present design is developed with relation to the short-haul plant.

I. INTRODUCTION

Since its manufacturing began in 1962, the T1 Carrier System[1] has been widely accepted by Bell System operating companies as an economic system of high performance for the transmission of voice signals in interoffice trunks. Although its use is growing and will continue to grow for voice transmission, it is significant that the T1 carrier regenerative repeatered line[2] with its 1.544 megabit per second transmission capability is introducing into the Bell System the first common carrier high-speed digital transmission facility. Its potential use as a facility for the growing market of wideband data transmission is apparent.

Because the T1 carrier system was primarily developed for interoffice trunks of the exchange and toll connecting type its repeatered line design has been optimized for comparatively short haul use from the standpoint of the economics of installation and maintenance. Although distances of 200 miles or more may be obtainable, repeatered line installations have generally not exceeded 50 miles. Thus, it may be expected that initial use of T1 carrier as a wideband data facility will be emphasized for two general applications:

(i) As an extension of interexchange wideband circuits into the exchange plant, and

(ii) In the provision of very limited networks for the transmission of very high speed data for certain special uses.

Wideband data covers a broad category of data signals requiring transmission bandwidths greater than those of voice facilities. With the exception of TV, program, and some special, limited government services, demands for wideband data services did not materially develop until about 1958. At this time requirements for government services expanded and an interest in commercial transmission of computer type of data developed.

To date, about one million equivalent voice channel miles of wideband services have been furnished by the Bell System and their signal formats have been of a wide variety, including analog video, two-level facsimile, synchronous and asynchronous serial data, and parallel tape-to-tape and tape-to-computer data. To meet this variety of requirements initially, a number of special terminals were developed on an accelerated design and manufacturing schedule. It was clear, however, that for a continued expansion of wideband services an organized arrangement of standard offerings must be developed.

A large number of the signals transmitted over the present facilities may be resolved into a class of two-level serial signals which include synchronous serial data, asynchronous data, and facsimile. The following equipment has been developed or is in the process of development to provide for these types of services at data rates corresponding to group band and to supergroup band transmission:

(i) Data sets for processing machine information into a standard baseband format.

(ii) Modems for N carrier, L carrier and T1 carrier to process these baseband signals into the carrier facilities.

(iii) Baseband repeater systems utilizing wire pairs for interconnecting the data set at the customer's location to the telephone office.

II. SYSTEM OBJECTIVES FOR T_1 CARRIER DATA TERMINALS

A prime objective in the design of data terminals for the T1 carrier system is that the line signal generated by these terminals be compatible with the existing T1 carrier regenerative repeatered line so that a line may be used interchangeably with a D1 bank signal or a data signal and that these signals may exist simultaneously in the same cable cross-section. Thus, consideration must be given to the pattern or format of the line signal generated by data terminals with at least two constraints. First, the character of the pattern on the line, including period and density, must not adversely affect other repeatered

lines through cable crosstalk; and second, the pattern must contain a sufficient number of pulses to maintain timing in the individual repeater clocks. Detailed requirements for the line signal as to level, impedance, bit rate and general format are well defined in the T1 carrier system.

Unfortunately, the requirements for the data signal to be transmitted cannot be as well defined. Even in respect to two-level serial signals, where some standardization is being attempted, the data rates are widely variable. In addition, many computer machines store and transfer data in the form of parallel words. In order to transpose these data signals into serial streams additional equipment is needed for "buffering" and word organization. It can be shown that in the case of T1 carrier it may be more efficient to inject these signals into the transmission system in the original parallel form, thus giving another general set of data signal requirements. Two specific designs of terminals will be described which accept some forms of data signals in these general classifications. In order to find general application, the terminals require a high degree of flexibility as to the form the input data signal can take.

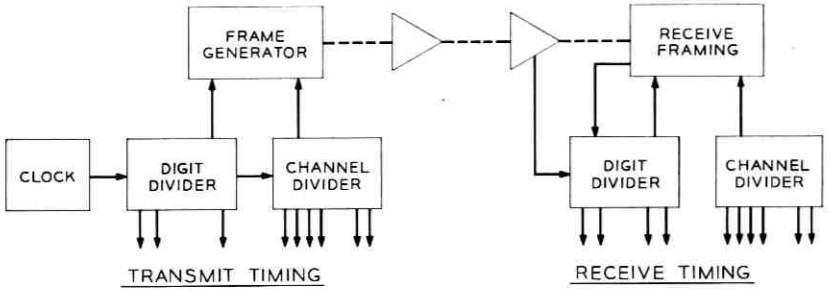
The terminals will also require the capability of time division multiplexing a number of data signals together on one T1 carrier line when each signal does not fully utilize the line capabilities. When the number of data signals of any one type or rate is not large, it may also be desirable to multiplex different types of data signals together or even data signals with voice signals from the D1 bank. The system requirements for these capabilities are considered in the design of the terminals.

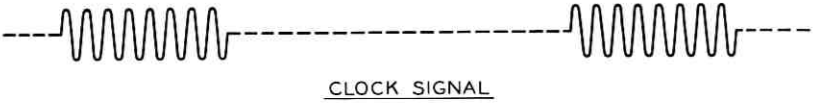
III. EQUIPMENT OBJECTIVES

Wideband data banks and modems are generally going to be installed in the telephone company central office. In addition, equipment designed to operate with a particular system, such as T1 carrier, should be located near that system's terminals. These premises imply that the equipment format should be chosen to be compatible with that of the "host" system.

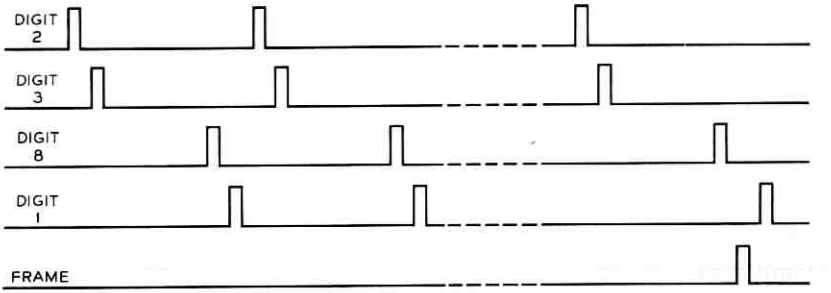
The circuits which are being designed for the T1 carrier wideband data banks and modems are fashioned after existing D1 bank circuits. As a matter of fact, many of the timing circuits for the data terminals are identical to those used in the D1 bank. These factors, coupled with others developed earlier in this section, resulted in an examination of the consequences of using D1 bank hardware for the data terminals.

There are several advantages in adopting the D1 bank equipment

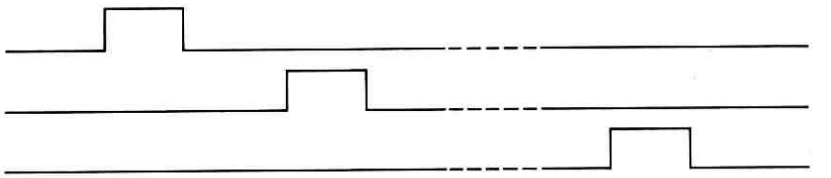




CLOCK SIGNAL



DIGIT TIMING SIGNALS



CHANNEL TIMING SIGNALS

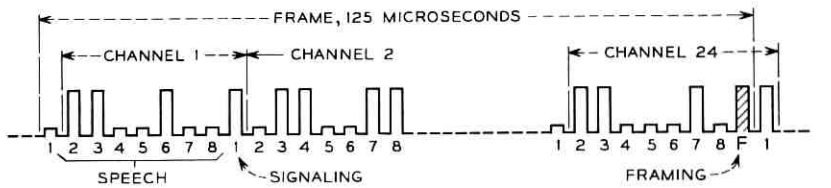


FIG. 1 — Block and timing diagrams for T1 carrier.

arrangements for data applications. Since, as already mentioned, the data terminals should be compatible with their voice terminal counterparts, the choice of existing D1 bank shelf castings and die-cast unit frames for data is a major step in achieving this compatibility. In addition, since the T1 carrier system is a high-production system, with some 130,000 group plug-in units shipped annually, distinct economic advantages are realizable by using the common piece parts and plug-in units when possible for the wideband data terminals. These considerations lead to the decision to use the same basic equipment design as that used in the T1 carrier D1 banks for the T1 carrier wideband data banks and modems.

Although most of the wideband data equipment could be in the central office, some of it will be installed on customer's premises, such as a computation center. In these cases, enclosures to house the normally relay-rack mounted equipment may be required to be complementary to the customer's installation.

IV. TERMINAL DESIGN

4.1 *Timing*

The basic control system of a time division multiplex terminal is its set of timing circuits. In order to derive the economic benefits of using standard D1 bank networks where applicable it is desirable that the timing circuit arrangements be similar. Fig. 1 shows functionally the basic components of these circuits. A crystal oscillator (1.544 mcps) provides the basic clock frequency and its output is used throughout the terminal for phase control of the signals. A digit generator normally divides the clock frequency by eight and provides 8 space and time separated outputs at this subdivided rate. A channel counter further subdivides the rate of the digit generator by 24, providing 24 space and time separated outputs, each at an 8-kc rate. For every 24th count of the channel counter (every 125 microseconds) the digit generator is controlled to count 9 instead of 8. This time slot contains a framing signal on the T1 carrier line whose function is to identify to the receiving timing circuits its unique position and thus the position of each of the 193 bits in every 125-microseconds frame period of the line signal. The receiver contains a clock which is controlled by the line signal to be synchronous with the oscillator in the transmitter. Dividers similar to those in the transmitting timing and framing logic circuitry provide a set of timing signals in the receiver identifiable with those generated in the transmitter.

In the D1 bank, the 24 voice input signals are each sampled in sequence every 125-microseconds frame period. These time divided samples are combined and each is sequentially coded into a 7-bit word by a common encoder circuit. An eighth bit is then added to each code word for signalling. It is natural, therefore, to find the eight bits of each sample grouped into words as shown in Fig. 1. It is neither natural nor convenient to force data signals into this format when time division multiplexing since it is desirable to provide in the bit stream equal sampling intervals for a channel. The periods of these intervals depend upon the number of T1 carrier line bits allocated to each channel. For example, assume a channel has been allocated one-half the repeatered line capability for transmission of its data signal. The timing for this channel can be controlled naturally from the timing derived from alternate line digits. Further, a channel requiring one eighth of the line can be controlled by a timing signal derived from every eighth digit. Since each of the 193 bits of a frame is uniquely identified, a variety of timing arrangements can be made using arrangements of the existing D1 bank circuitry.

4.2 *Serial Data Terminal*

One important application of data on T1 carrier will be as an extension of wideband toll data circuits through the exchange and local loop plants. For this a very flexible terminal is required to handle a variety of types of signals. Even if limited to two-level serial data signals this variety includes synchronous serial data of various rates, asynchronous serial data, and two-level facsimile. These signals may be considered as a class whose transitions between two levels occur at random with the minimum interval between transitions limited by the maximum data rate. The essential information to be transmitted for these signals is the time of transition and the state (1 or 0) after the transition.

Because the T1 carrier line is a synchronous digital facility, the analog information of transition time must be quantized and encoded in some manner. In the terminal design developed for this application these functions are accomplished efficiently with comparatively simple circuitry, meeting a maximum quantizing error objective of less than ± 10 per cent of the minimum data bit interval.

The principle of the "sliding index" provides the basis for the method of timing encoding in this terminal as well as the parallel data terminal to be described later. This principle was originally applied by Messrs. C. G. Davis and L. C. Thomas in the design of a parallel tape-to-computer data terminal for a dedicated T1 carrier line.[3] The essen-

tial feature of the "sliding index" approach is that it allows the insertion of a data word into a bit stream with the first available bit in the stream rather than holding the data for a particular position in the frame of the line signal. The details of the process will become more apparent in the following descriptions.

Consider first the simple functional block and timing diagram of Fig. 2. Line A shows a timing signal which is derived from clock and timing circuitry. As described earlier, this marks the time slots available in the T1 carrier line signal for one of a number of multiplexed data signals and may be every 2nd or every 8th or every *n*th bit as selected. Line B shows the data signal as applied to the terminal coder with transitions randomly distributed among the timing periods. In the absence of transitions the coder output is zero in the selected time slots (line C). When a transition occurs the next available time slot is indexed with a *ONE*, transmitting to the receiver decoder information that a transition has occurred between this time slot and the last preceding one. The coder then marks the next successive time slot with a *ONE* if the transition occurred in the early half of this period or a *ZERO* if it occurred in the late half. The third successive time slot is

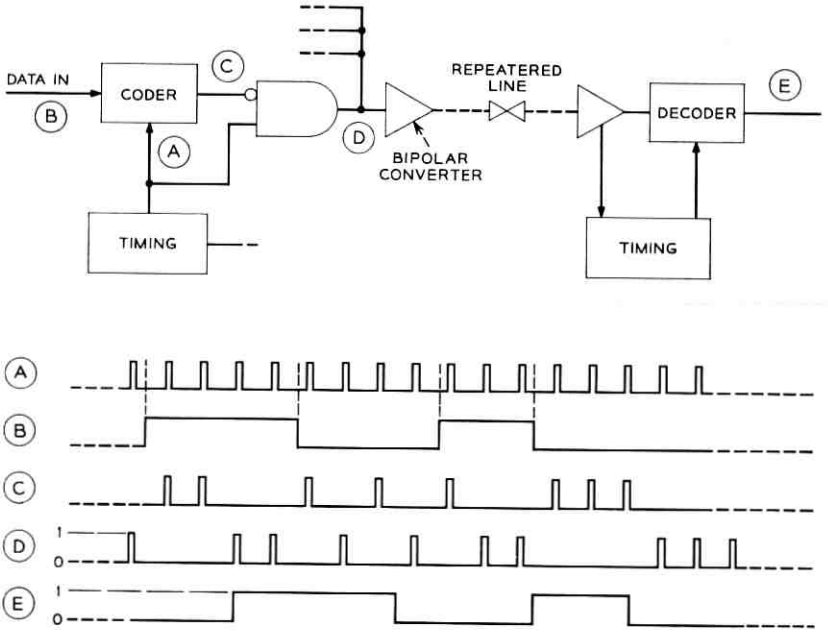


FIG. 2 — Functional block and timing diagram for the serial data terminal logic.

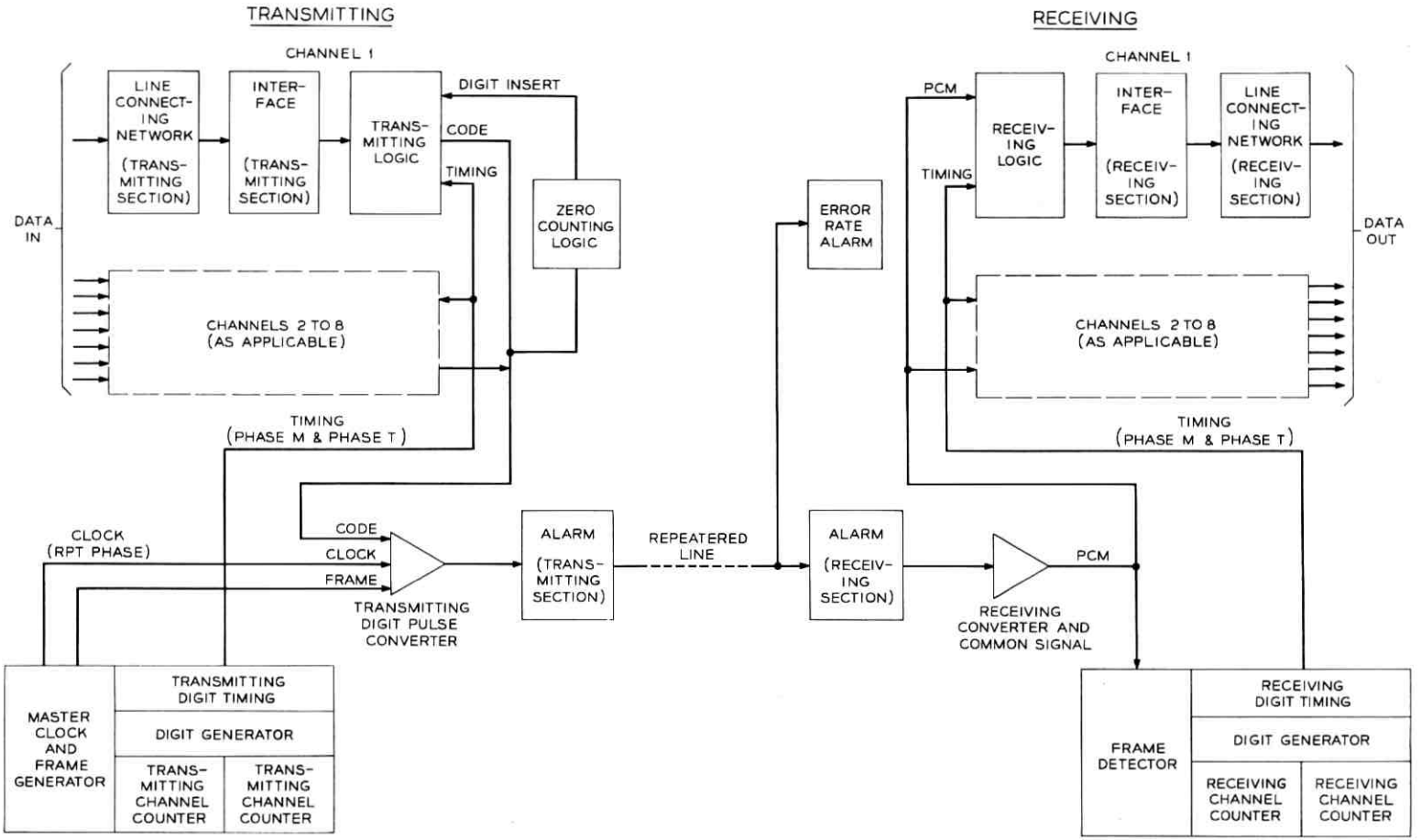


FIG. 3 — Block diagram of T1WB-1.

marked with a *ZERO* for a data transition from *ZERO* to *ONE* or a *ONE* for a transition from *ONE* to *ZERO*.

It should be noted here that if this output signal is logically inverted, then *ONES* will be transmitted in the absence of transitions. In general, this inversion is made because it provides a better timing signal on the line for the repeater clock circuits (line D).

The receiver decoder logic resolves this information into a data transition occurring within one-half of the selected timing interval plus a fixed delay. Since the minimum interval between data transitions must be limited to three selected T1 carrier timing intervals, and the transition is quantized to within $\frac{1}{2}$ this selected timing interval, the quantizing error is a maximum of $\pm\frac{1}{12}$ or ± 8.3 per cent, of this minimum data transition interval.

The foregoing is the process which is the basis for a set of data channel banks for asynchronous serial data transmission over T1 carrier lines. A number of functions must be performed in the terminals in addition to the coding process. Fig. 3 shows a block diagram of a data channel bank for up to 8 independent data signal inputs and includes the circuits for these auxiliary functions.

The arrangement of the timing circuits and their functions have been described in the general section on timing. In this 8-channel circuit specifically a digit timing unit is provided which steers the proper timing signals to each of the channel logic circuits such that the coded data signal for Channel 1 is transmitted in the time slot of Digit 1 on the line (Fig. 1), the data signal for Channel 2 in the time slot of Digit 2 and so forth.

In the coding process described earlier the data input was idealized as a two-level signal. The data signal received by the terminal as part of the standard wideband network will require amplitude regeneration to obtain this signal for several reasons. First, the data set processes the two-level signal as received from the customer's machine by removing low frequency energy from the signal in a controlled manner.[4] This is done to facilitate transmission over the analog facilities. In addition, the signal spectrum may be band limited and noise may be added in these facilities. The purpose of the transmitting Line Connecting Network and the Interface Network is to regenerate this signal to the two-level format. In some analog facilities, such as those including L-Multiplex, the frequency band for the data signal is limited to less than the bit rate frequency. For optimum shaping of the data signal prior to detection, a Line Connecting Network is provided which includes a network to roll off the band to a frequency 50 per cent above the signalling rate

or to $\frac{3}{4}$ the bit rate frequency. When the intervening facilities do not restrict the band this severely, as in the case of baseband repeatered lines, a second optional network is provided which rolls off the band to 100 per cent above the signalling rate or to the bit rate frequency. These networks employ transfer functions which satisfy Nyquist's criteria for periodic zero crossings in the pulse response. When the data terminal is located near the data set such that little noise is added to the signal no shaping or roll-off network is required.

The Interface Network detects and regenerates this signal by first restoring its dc and low frequency components by means of quantized feedback, then slicing it. At the receiving end the Interface and Line Connecting Networks remove the low frequency energy from the signal to the same extent as the data set, thus preparing it for transmission through other facilities.

The combined outputs of the logic circuits consist of a stream of unipolar pulses. These are applied to a Digit Pulse Converter whose function is to alternate the pulses from plus to minus creating the bipolar pulse stream required for the T1 carrier line. The receiving converter derives the basic clock signal from this pulse stream and converts the received signal back to its unipolar format. This unipolar signal is applied to all receiving logic circuits where it is demultiplexed by the timing signals and decoded.

Two alarm arrangements are provided. One circuit presents an alarm when framing is lost in the receiving bank. Its features are similar to those provided in the D1 bank, including the capability of terminal looping. However, the T1 carrier framing circuits are comparatively rugged and require error rates in the order of 10^{-3} in order to initiate an alarm. For this reason a second alarm circuit is provided which alarms directly on line error rate. This circuit includes a bipolar violation detector which provides a measure of line errors. By means of an integrating circuit an alarm indication is obtained when the error rate exceeds a predetermined value, say 10^{-6}.

Due to the inherent flexibility of time division multiplex systems, only minor changes in timing arrangements are required to obtain a variety of data speed options. For example, if the timing signal is arranged to code an asynchronous data signal into every eighth T1 carrier bit, this channel will have a maximum data rate capability of 64 kilobits per second, adequate for the 50-kilobit signals to be transmitted in the group band over L-multiplex toll facilities. Eight such signals can be multiplexed on one T1 carrier line. Further, data rates up to 256 kilobits per second may be transmitted over a channel whose

signals are coded into every second T1 carrier bit, providing a capability corresponding to supergroup transmission in L Multiplex. Table I lists some typical asynchronous serial data rates with their limits of timing error due to quantizing.

The framing time slot in the line signal introduces an additional timing error which is not included in Table I. This error occurs when a transition is such that its 3-bit code sequence encompasses the framing time slot; it introduces an additional timing delay of 0.6 microseconds. This is significant for the case of 256 kilobits/second data transmission, but only occurs on about 2 per cent of the transitions. A number of subjective tests have been made transmitting facsimile copy at this rate. No apparent degradation in performance occurred which could be attributed to the framing error. If further tests and field experience show this error to be of some importance additional logic circuitry may be added to constrain this error to occur only when the data transition occurs during the framing time slot, or about 0.5 per cent of the transitions.

4.3 *T1WB-1 and T1WB-2 Wideband Bank Equipment Design*

The first standard offerings for wideband data terminals for T1 carrier, based upon the foregoing discussion, are the T1WB-1 and T1WB-2 wideband banks. The banks are made up of die-cast aluminum shelves which mount the plug-in timing, logic, line connecting and interface units, and fabricated panels which contain other line connecting and miscellaneous circuits. Figs. 4 and 5 illustrate the use of T1 carrier hardware in these banks.

The two banks described here differ only in their respective data channel capacities. The T1WB-1 wideband bank is arranged to transmit and receive up to eight two-level asynchronous serial data or facsimile signals over a T1 carrier repeatered line. The T1WB-2 wideband bank is a scaled-down version, being arranged to process up to two 250-kilobit signals.

TABLE I

T1 Line Loading	Number of Channels Per Line	Max. Data Rate In kb Per Sec.	Max. Timing Error
1/8	8	64	±1.3 Microseconds
1/4	4	128	±0.65 Microseconds
1/2	2	256	±0.35 Microseconds
1/1	1	512	±0.17 Microseconds

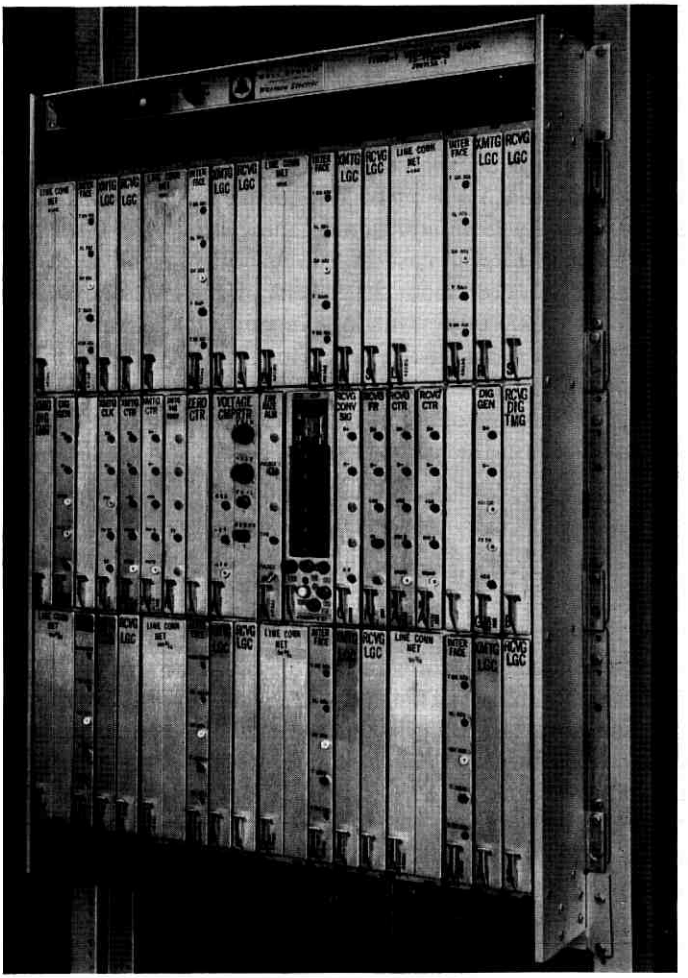


FIG. 4 — T1WB-1.

Hardware in the banks has been grouped into two categories. The equipment for processing the customer's baseband signal into a PCM signal is termed channel equipment. This equipment can be plugged into the terminal on a per customer basis. Equipment which is shared by all channels to derive timing information and to provide fuse and alarm indications is termed common equipment. A portion of this latter equipment consists of panel mounted jacks, battery filters and terminal strips for access to the banks. The remainder, the active circuitry, is

provided as plug-in units. The two banks make use of similar configurations of these equipments as may be seen in Figs. 4 and 5.

The T1WB-1 wideband bank occupies a space of nineteen $1\frac{3}{4}$-inch by 23-inch mounting plates and weighs 130 pounds when fully equipped. The T1WB-2 occupies a space of thirteen $1\frac{3}{4}$-inch by 23-inch mounting plates and weighs 80 pounds when fully equipped. The banks may be mounted on 10-inch deep bulb angle bay frameworks or 12-inch deep cable duct bay frameworks. All installer wiring is brought to terminal strips on the rear of the banks. This, of course, means that no back-to-back bay lineups are possible for T1 carrier wideband data installations. However, the same restriction exists for the T1 carrier voice terminal.

Since both banks are functionally and mechanically similar except as outlined above, further description will not be identified with either bank. Unless otherwise stated, comments made will apply to both.

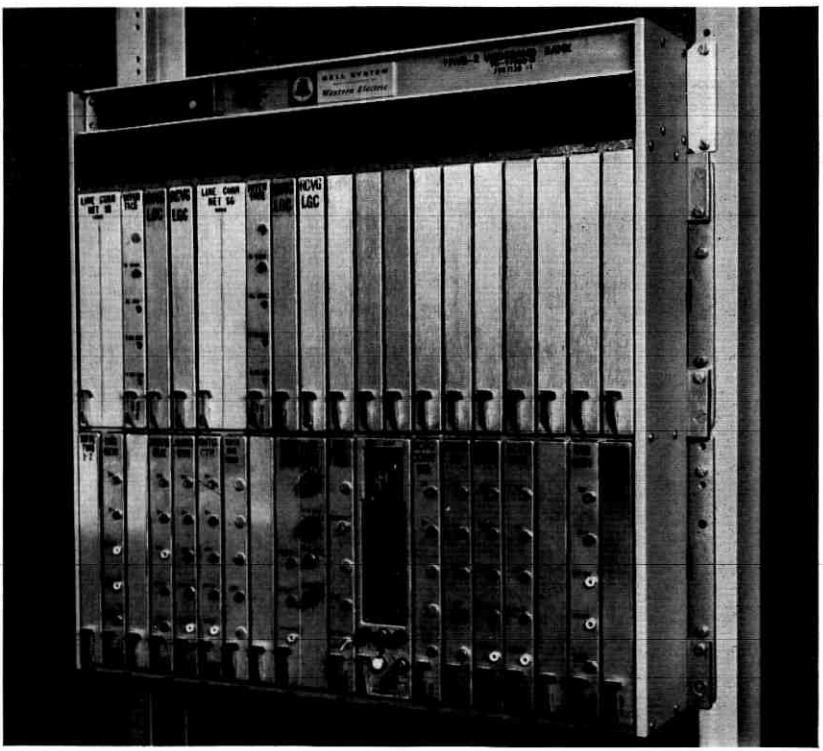


FIG. 5 — T1WB-2.

The timing and channel shelf assemblies consist of die-cast aluminum piece parts bolted together in such a way as to accept the plug-in units. Fig. 6 shows an assembly of two shelves. It may be noted that the assembly consists of two connector mounting die castings and three other die castings which contain the tracks and retaining slots into which the plug-in units slide and lock in place. The entire assembly is strengthened by dividers which also serve as interposition shields. The connector mounting die casting is designed to hold up to twenty pairs of connectors, each pair mounted one vertically over the other as shown. This connector arrangement provides a maximum of forty-two terminals for each plug-in unit. One shelf assembly can be equipped with twenty "single" plug-in units or ten "double" plug-in units or combinations of the two types. The connector-mounting die casting has mounting holes on the rear for attaching terminal strips.

Typical plug-in units are shown in Fig. 7, one the "single" type and the other the "double" type. Either unit is made up of a die-cast aluminum frame, a printed wiring board and one or two plugs. The

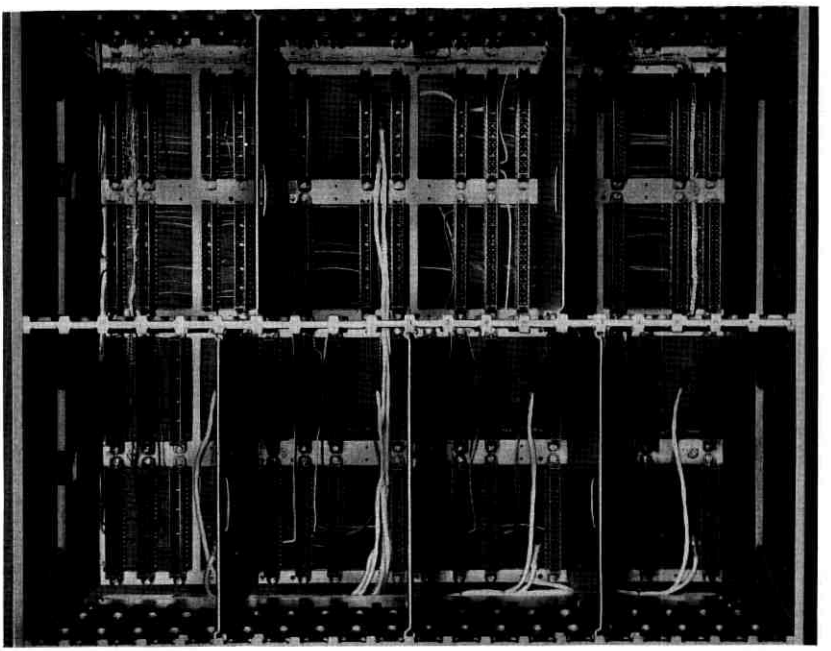


FIG. 6 — Typical shelf assembly showing type of construction and placement of connectors.

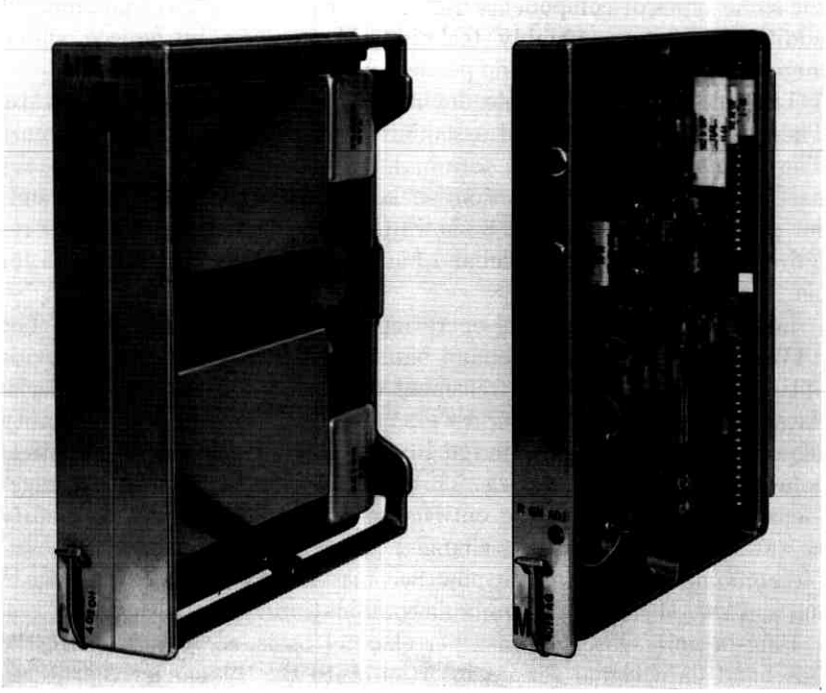


FIG. 7 — Typical T1WB-1 or T1WB-2 plug-in units.

frame has a latch mechanism not unlike others used in the telephone plant which locks the unit in place when inserted into the shelf assembly. The latch also serves to break the connector contact pressure when the unit is to be extracted. The printed wiring board is either of a fire retardent phenol fiber material or an epoxy glass material depending upon the mass of the component apparatus to be mounted upon it. It may be double-sided or single-sided depending upon the complexity of the circuit. In the double-sided case, connections between sides of the board are provided by through-straps. The over-all dimensions of the single plug-in units are approximately $8\frac{5}{8}$ inches by $7\frac{29}{32}$ inches by 1 inch, exclusive of the latch and plug. The double plug-in units are approximately $8\frac{5}{8}$ inches by $7\frac{29}{32}$ inches by 2 inches, exclusive of the latch and plug.

Out of 27 separate codes of plug-in units required for all modes of operation of the T1WB-1 and T1WB-2 wideband banks, 7 are already used in the T1 carrier system. The 20 new codes were designed using

the same types of components used in T1 carrier whenever possible. In addition, no new assembly techniques have been introduced which might incur retraining of shop personnel.

The line connector panels deviate from the T1 carrier hardware. These panels are fabricated assemblies of 0.090-inch sheet aluminum. They mount power filters, terminal strips and access jacks for the baseband channels. In addition, sockets are mounted on these panels for plug-in span pads. Fig. 8 shows the two line connector panels required for the T1WB-1 wideband bank. The top panel is also used for the T1WB-2.

Let us now explore the operational features of the banks. The T1WB-1 and T1WB-2 wideband banks have been designed to provide optimum flexibility in the treatment of two-level asynchronous serial data or facsimile signals over a wide range of data rates. The flexibility required was attained by design in the common circuits for practical combinations of these signals. This resulted in an equipment arrangement unhindered by wiring options necessary for the field to modify as data signal requirements changed. Administering the various possible combinations of signals now becomes simply a choice of plug-in units. Table II shows the available combinations.

Plug-in units, inserted on a per channel basis, serve to prepare the baseband data signal for its insertion onto the T1 carrier repeatered line. Different data rates and the necessary baseband line equalization are achieved by selecting the proper plug-in units. In addition, other

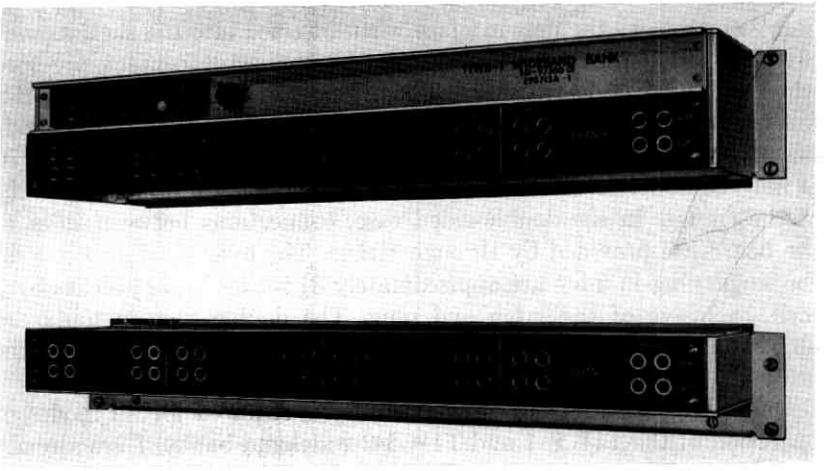


FIG. 8 — Line connector panels.

TABLE II

Wideband Bank	Multiplex
T1WB-1	8 channel — 50 kb
	4 channel — 50 kb 1 channel — 250 kb
	2 channel — 250 kb
T1WB-2	2 channel — 250 kb

units are available to limit the number of successive *ZEROS* in the line signal for satisfactory repeater operation. Timing for the data bank is provided by another group of plug-in units.

The T1WB-1 and T1WB-2 wideband banks are equipped with optional pads or equalizers to compensate for the cable length between the bank and the T1 carrier office repeater. This is a wired option, since it is considered unlikely that, once installed, it will have to be changed.

If the data bank is installed in a location having no other T1 carrier facilities, auxiliary panels are provided. These panels mount the necessary equipment for connecting the bank to a repeatered line and for fault locating on that line. Panel designs are available in the T1 carrier system to perform these functions.

The T1 carrier system may provide a very convenient means for interconnection of central office equipment and a data set at a customer's location even though the repeatered line is dedicated to a single data signal. For these applications, the T1WM-1 wideband modem has been developed which is basically a simplified version of the T1WB-1 and -2 banks. It consists of a single shelf of plug-in networks similar to those used in the banks. However, most of the timing units required for multiplexing and framing are eliminated. This modem is capable of transmitting serial data signals up to 500-kilobit rates.

4.4 *Terminals for Parallel Data*

The geographical diversification of large research and development organizations with a common need for high-speed computer facilities creates a requirement for data transmission at rates above those generally required for business machine data or facsimile. One example of such a complex is represented by three major Bell Telephone Laboratories locations at Murray Hill, Holmdel and Whippany, New Jersey. Each of these locations contains a computation center equipped with

large, high-speed computer facilities. The need for load sharing, during heavy load periods or computer "down time" at one location, is but one reason for interconnecting these centers with data transmission facilities. Because of its obvious convenience for experimentation and testing, this Laboratories' computer complex was selected as a model for the development of experimental T1 carrier data terminals for this general type of service. It will be shown, however, that the basic design contains sufficient flexibility to operate with a large variety of computer systems.

The interchange of information in the present Bell Laboratories' computer complex principally involves tape machine to tape machine or tape machine to computer data transmission. Fig. 9 shows the essential units for this operation when the transmitting tape machine is remotely located. Under the control of slow speed signals sent from the receiving equipment the tape machine reads data from the tape, transmitting it to the data terminal in the form of groups of parallel data characters. The data terminal puts this information into a form suitable for transmission over a T1 carrier line and restores it to the original parallel character format at the receiving end.

The character size, that is the number of bits per character, and the character rate are dependent upon a particular tape machine design. In the Bell Laboratories service tape machines are used which transmit a 7-bit (or 7-level) character at nominal rates up to 90 kilocharacters per second. Because of the nature of tape drive mechanisms this character rate is somewhat "elastic" and may have momentary variations as great as ±20 per cent of the nominal rate. It is important, however, that the character spacing be preserved in transmission within some

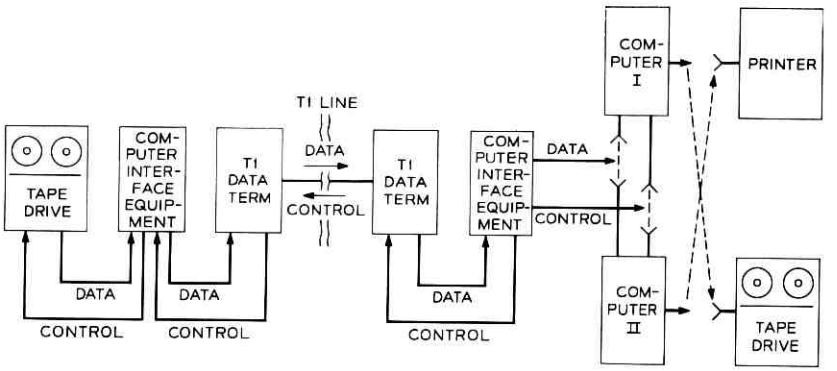


FIG. 9 — Remote computer control operation block diagram.

limitation since this spacing may contain information essential to receiving computer equipment. Since much of the character spacing tolerance is used up in tape speed variation, little is left for the transmission system. The principal point to be made here is that T1 carrier data terminals must be capable of transmitting these signals at asynchronous character rates while quantizing the character timing information within about ±15 per cent of the minimum character period.

In general, the prior discussion of terminal timing control for serial data terminals is also applicable to the parallel data terminal. However, because of the data rates involved, the terminal has been designed principally for a dedicated data service with the multiplexing capability used for auxiliary signals such as the slow speed control signals. The block diagram of Fig. 10 shows circuits which generate a set of timing signals including framing. Of this set of signals the data clock signal controls the timing of the input data signals and marks the bit intervals on the line in which the data signals are transmitted. As shown in the timing diagram, Fig. 11, these bit intervals consist of all but the framing time slot and those selected for the auxiliary control information. In the absence of data input signals *ONES* are transmitted on the T1 carrier line in these time slots.

The data characters are transmitted to the terminal from the computer equipment on seven parallel leads, each lead handling one bit of the character in a bipolar format of plus and minus pulses for *ONES* and no signal for *ZEROS*. A data character is recognized when a pulse (a *ONE*) is transmitted on any one or more leads. On this event all leads are read and those with no pulses are considered to be transmitting *ZEROS*. No significance exists in the polarity of the pulses nor does any relation exist in the polarity of a pulse on one lead with respect to any other.

When a data character is transmitted to the terminal the pulses are rectified to a unipolar format and detected in the interface circuit. The detected signals are applied to respective stages of the register and the data character is stored there. The occurrence of the data character is also recognized by the data logic circuit which, on this occurrence, applies the data clock signal to the shift control of the register. This causes first a *ZERO* index to be transmitted on the line, followed immediately by the seven bits of the data character. Should a control or framing time slot occur during this interval, the shifting sequence will stop for that time slot, resuming its count after the auxiliary signal is transmitted.

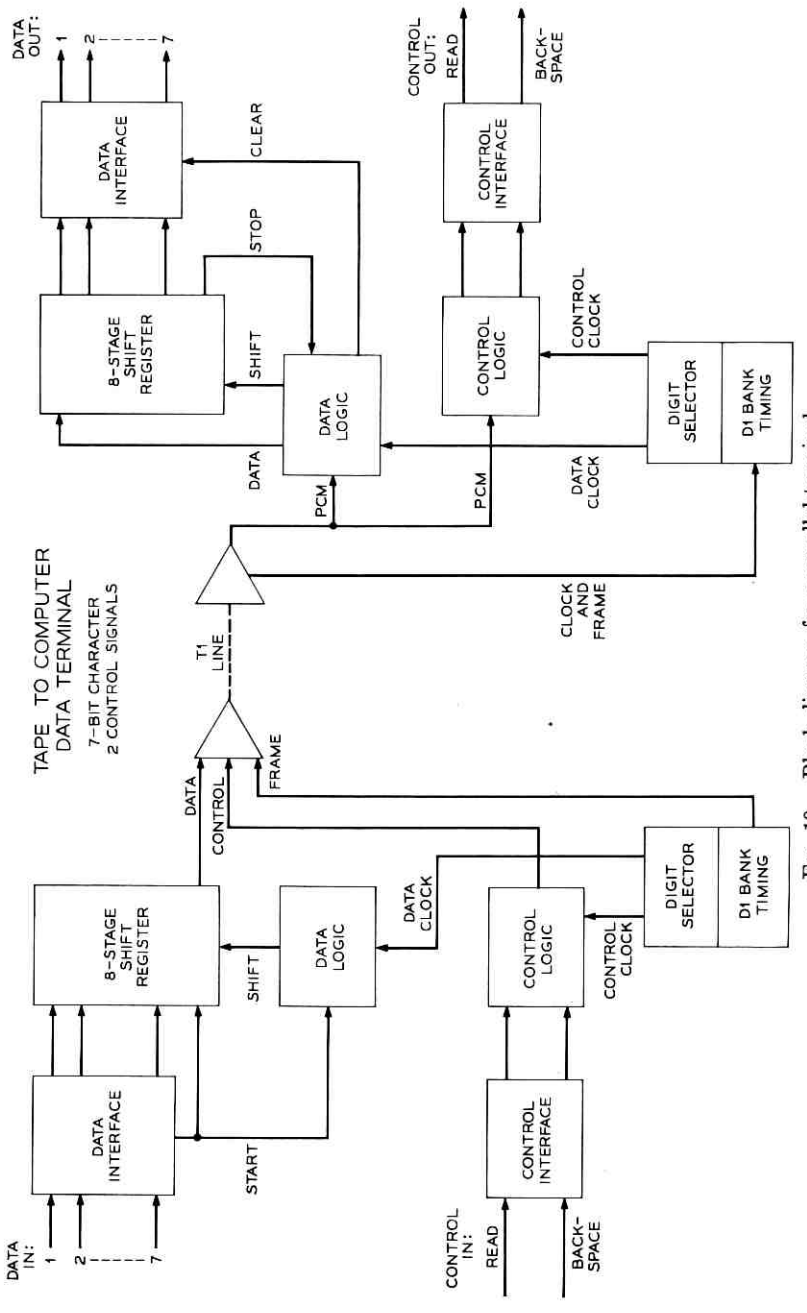


FIG. 10 — Block diagram for a parallel terminal.

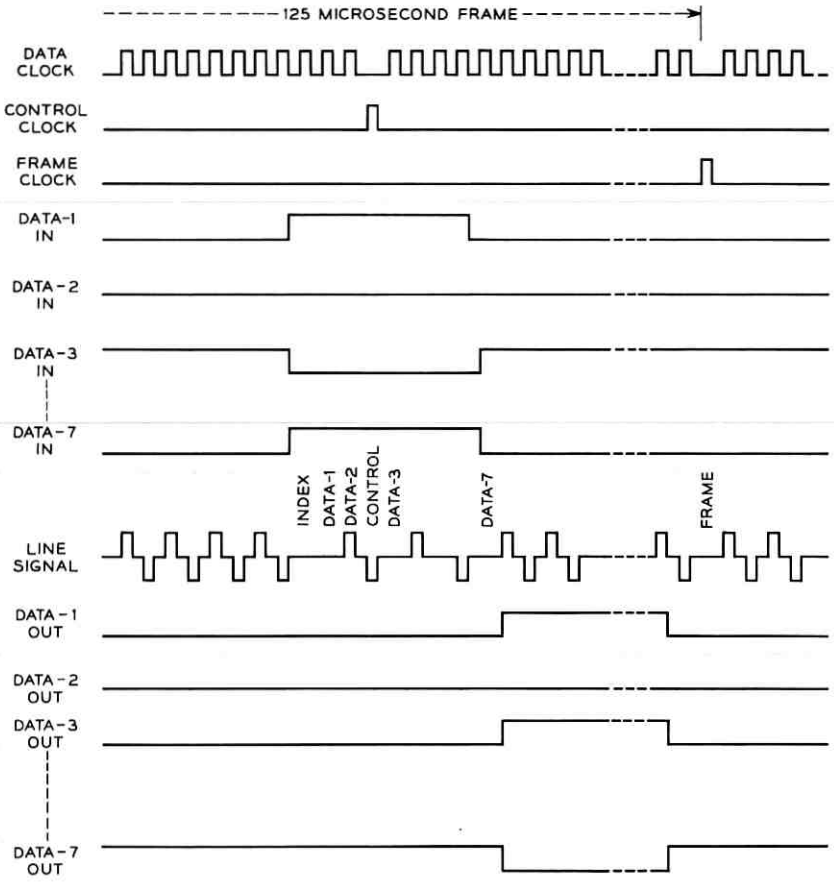


FIG. 11 — Timing diagram for a parallel terminal.

In the receiver, the *ZERO* index is recognized as the address of the character by the receiving data logic. This circuit then applies a data clock signal to the shift control of the register, causing the next seven data bits to be shifted into the register. This data clock signal is identical to and synchronized with that of the transmitter, and thus excludes the timing interval during which framing or control signals occur. When a control or framing bit occurs within a data word the shifting sequence stops for that interval allowing only bits of the data word to enter the register. At the end of the count, the shifting sequence is stopped and the data character is transferred in parallel form to the interface. The interface generates bipolar pulses for *ONES* in the character, ap-

plying them to the parallel leads for transmission to the computer equipment.

What has been shown so far is only half of the full-duplex capabilities of this data system. With both transmitting and receiving terminals at each end another set of independent parallel data signals may be sent in the opposite direction. Multiplexed with this set, however, are the slow speed control signals, alluded to earlier, for the first set of parallel data signals. These control signals are simple instructions or commands to the sending computer equipment and consist of binary impulses spaced not less than 6 to 8 milliseconds apart. A separate lead is supplied in the receiving computer equipment for each control signal required.

The coding and multiplexing of these control signals is rather easily implemented. For each control bit required, a T1 carrier bit is selected by the timing control from every 193-bit frame. The timing is so arranged that the control bits are approximately evenly distributed in the frame. In the absence of control data, *ONES* are transmitted on the line in these bit periods. When a control impulse is applied on one of the control inputs, *ZEROS* are transmitted in its particular bit period for three successive frames. Majority logic circuits in the receiver recognize this signal on a 2 out of 3 bit basis, providing for single bit error correction. From this information an impulse is regenerated and applied on the appropriate control lead to the computer equipment. Control impulses as closely spaced as 500 microseconds may be accommodated.

From the foregoing it may be seen readily that the character timing of the high speed data is preserved within the quantizing of the data timing signal. Except when the character is transmitted over a framing or control bit period the resulting distortion is within ± 0.33 microseconds. The framing or control bit will add 0.65 microseconds delay to characters encompassing their time slots. It may also be seen that the character size is not dependent upon the basic terminal design. By adding or removing certain stages of the terminal the character size capability may be increased or decreased with an inversely proportionate transmission rate capability. This is shown in the following relationship and summarized for some typical character sizes in Table III.

$$\text{Max. Char. Rate} = \frac{1544 - 8(C + 1)}{N + 1} \text{ kilocharacters per second}$$

where N is the number of bits per character, and C is the number of control channels.

TABLE III

Bits per Character	No. of Control Channels	Max. Character Rate	Nominal Timing Error in % of Min. Char. Interval*
12	4	115 kc	±3.8%
8	2	168 kc	±5.6%
4	2	304 kc	±10.0%

* Delay of framing or control bit is not included.

4.5 *Parallel Data Terminal Equipment*

A parallel data terminal has been implemented to process data characters up to 12 bits in length and as many as 6 control signals (see Fig. 12). This parallel data terminal consists of signal transforming equipment and common timing and alarm equipment. The modem occupies a space of eighteen $1\frac{3}{4}$-inch by 23-inch mounting plates and weighs approximately 130 pounds when fully equipped. It may be mounted on a 10-inch deep bulb angle bay framework or a 12-inch deep cable duct bay framework. As in the other T1 carrier data terminals, all the installer wiring is brought to terminal strips on the rear of the unit.

Continuing the philosophy that once the unit is installed, no further installer effort is required to make changes on it, the parallel terminal provides optional features on either a plug-in basis or a solderless wrap strap basis. Fig. 13 shows a rear view of the modem in which may be seen these strap options.

The construction of the shelves and panels is identical to that already described for the T1WB-1 and T1WB-2. Die-cast aluminum shelves and fabricated aluminum panels are the principal structural members. The plug-in units are single-sided and double-sided printed wiring boards mounted in die-cast aluminum unit frames. Fig. 14 shows a typical plug-in unit. This terminal makes use of "cord wood" component packages, as shown in Fig. 15, to derive a more efficient packing factor in its plug-in units. In the illustration, the "cord wood" module is a flip-flop having sufficient output leads to perform a variety of logic functions (see Fig. 16).

Although the modular concept is generally more expensive than the laying down of components directly to the printed wiring board, its use in this terminal has resulted in at least a 25 per cent reduction in total terminal volume. Conventional component placement would have required more plug-in units and, hence, another shelf to mount them.

This parallel data terminal provides flexibility of word size and the number of control channels in a unique manner. Fig. 17 illustrates how

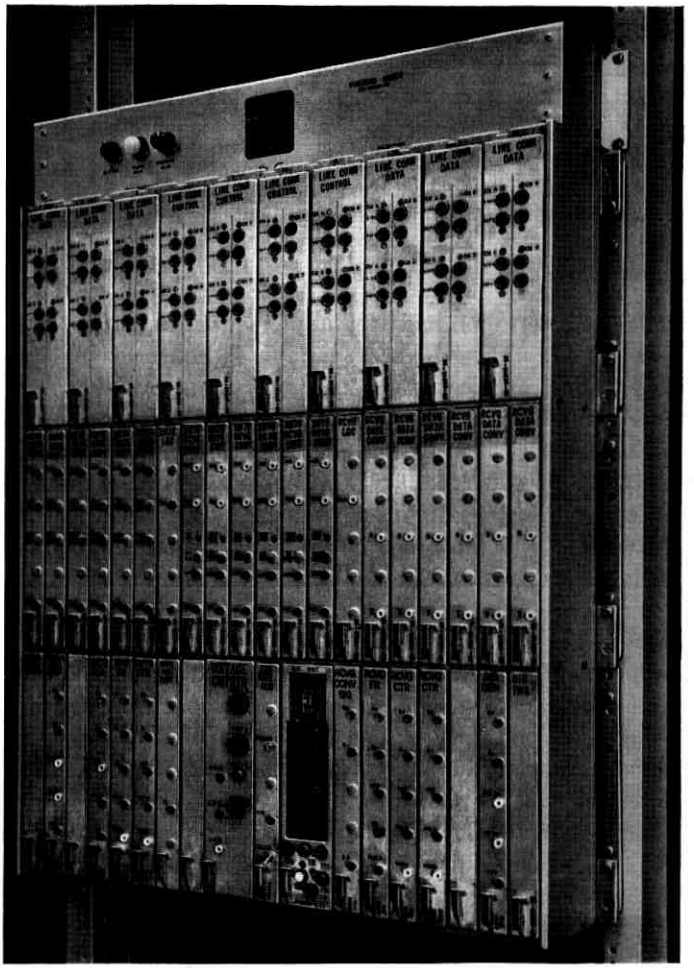


FIG. 12 — Parallel data terminal.

the transmitting side of the terminal is arranged for processing a 6- or an 8-bit character.

Plug-in units called Data Converters are each equipped with two separate interface circuits (I) and shift register stages (R). The shelf wiring is arranged to provide the unit interconnection as shown in the figure. A second type of plug-in unit called the Data Logic Unit is back wired to Data Converter 1.

As developed earlier, a pulse or *ONE* on any of the "bit" leads is recognized by the Data Logic unit via the "sense" lead and the register is read out serially.

The important thing to note here is that by plugging the proper number of data converters into the shelf and inserting the Data Logic unit into the slot next to the first Data Converter, any number of bits up to the capacity of the terminal may be accommodated. A strap option in the back wiring further provides for the handling of even or odd bit characters.

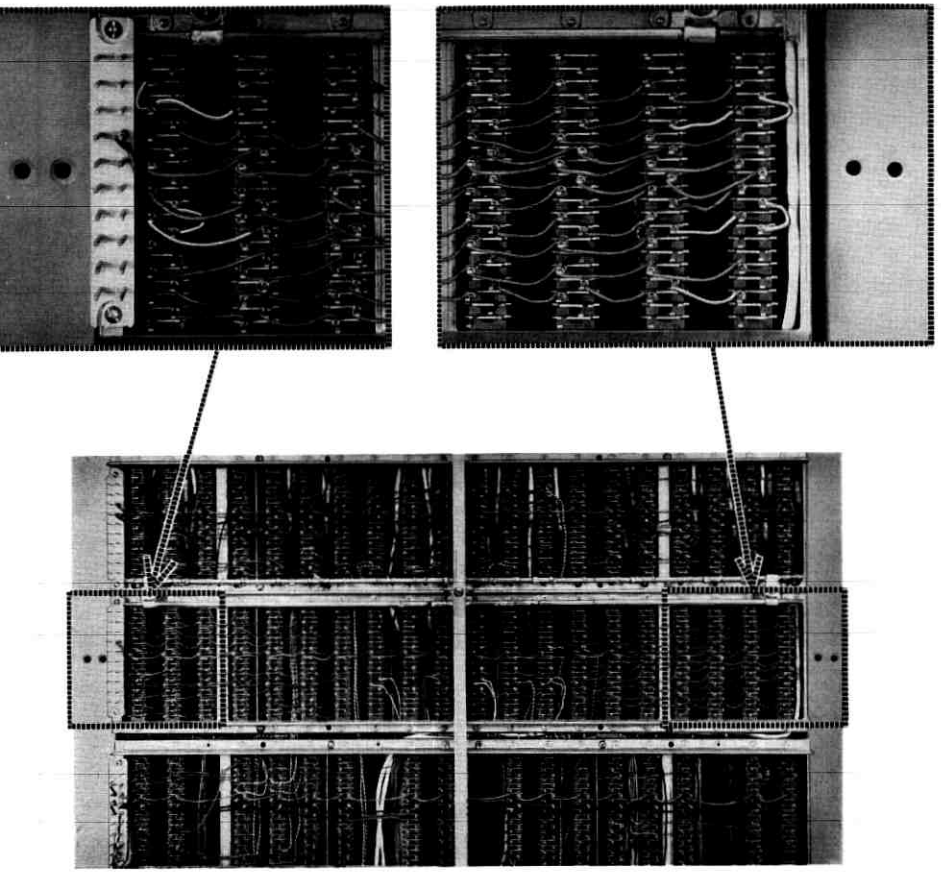


FIG. 13 — Rear view of parallel data terminal showing specifically the location of the strap options.

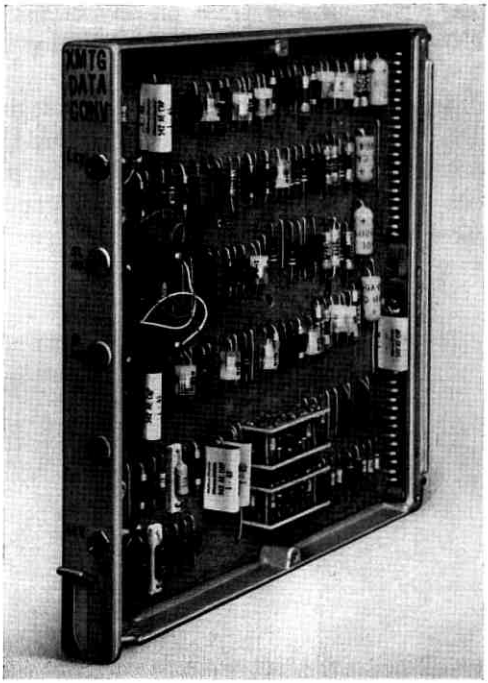


FIG. 14 — Typical parallel data terminal plug-in unit with "cord wood" modules in place.

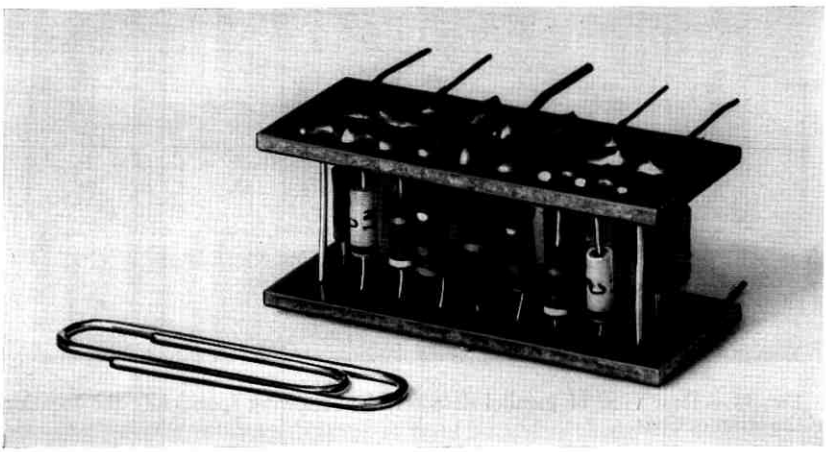


FIG. 15 — Flip-flop "cord wood" module.

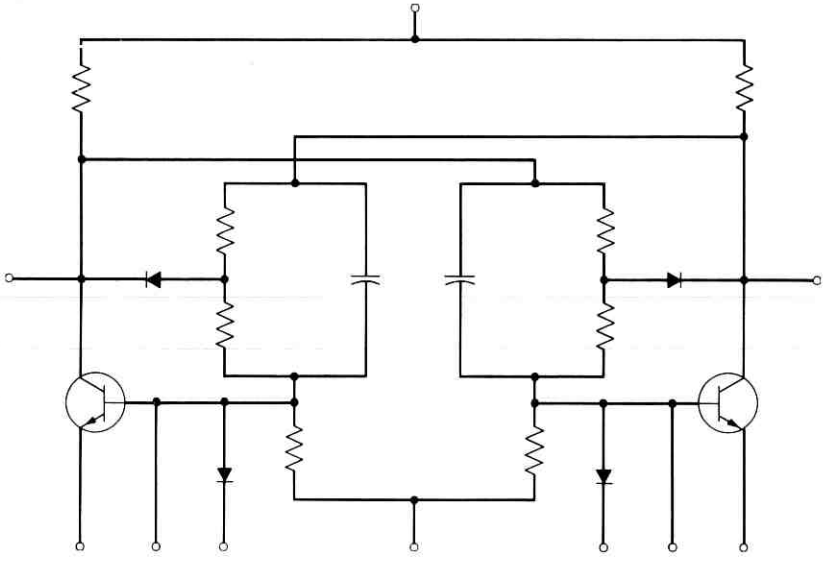


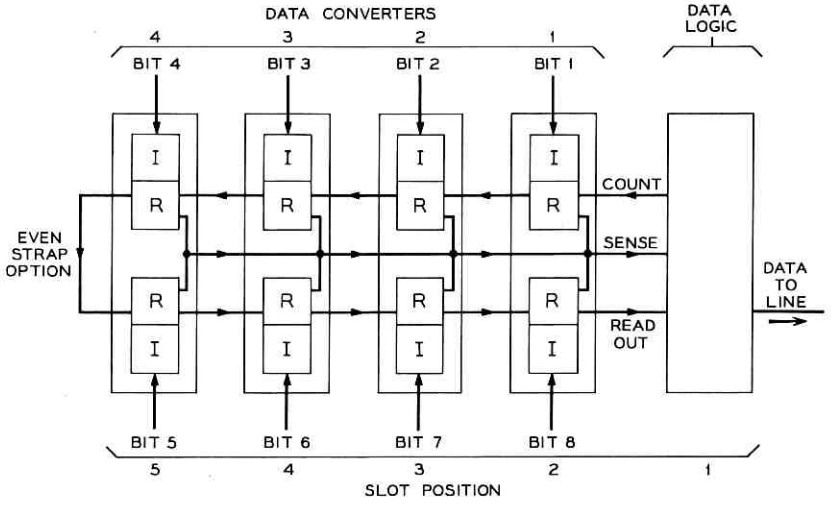
FIG. 16 — Schematic of flip-flop circuit which is packaged as "cord wood" module.

The receiving side of the terminal for data character processing uses the same mechanical implementation as does also the control channel circuitry.

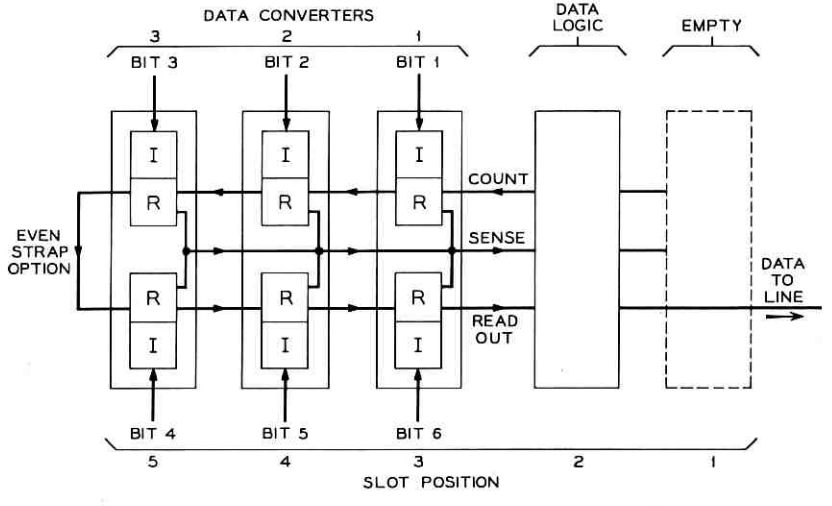
Fig. 18 shows the terminal arranged to transmit an 8-bit character and a 6-bit character plus control signals. Note the differences in the arrangement of plug-in converters and logic boards.

4.6 *Field Performance*

Since June of 1962, T1 carrier has provided for data transmission among the Bell Laboratories computation centers at Holmdel, Whippany and Murray Hill. This service requires transmission of the data over the distances of roughly 20, 40 and 60 miles. The system made possible the remote operation of the computers through the transmission of 7-level tape data at nominal speeds of 62.5 kilocharacters per second. It has been used primarily for load sharing among the locations. In the spring of 1963, laboratory models of terminals based on the design described above replaced earlier terminal equipment. After a period of shakedown of the terminals and the T1 carrier repeatered lines, the system reached and sustained reliable performance at very low error rates. Record retransmissions which result from transmission errors, have averaged about 0.01 per cent for 1000 character records, corresponding to an average T1 carrier line error rate of 10^{-8}. This per-



(a) EIGHT-BIT CHARACTER



(b) SIX-BIT CHARACTER

FIG. 17 — Block diagram of plug-in arrangement for data transmitting side.

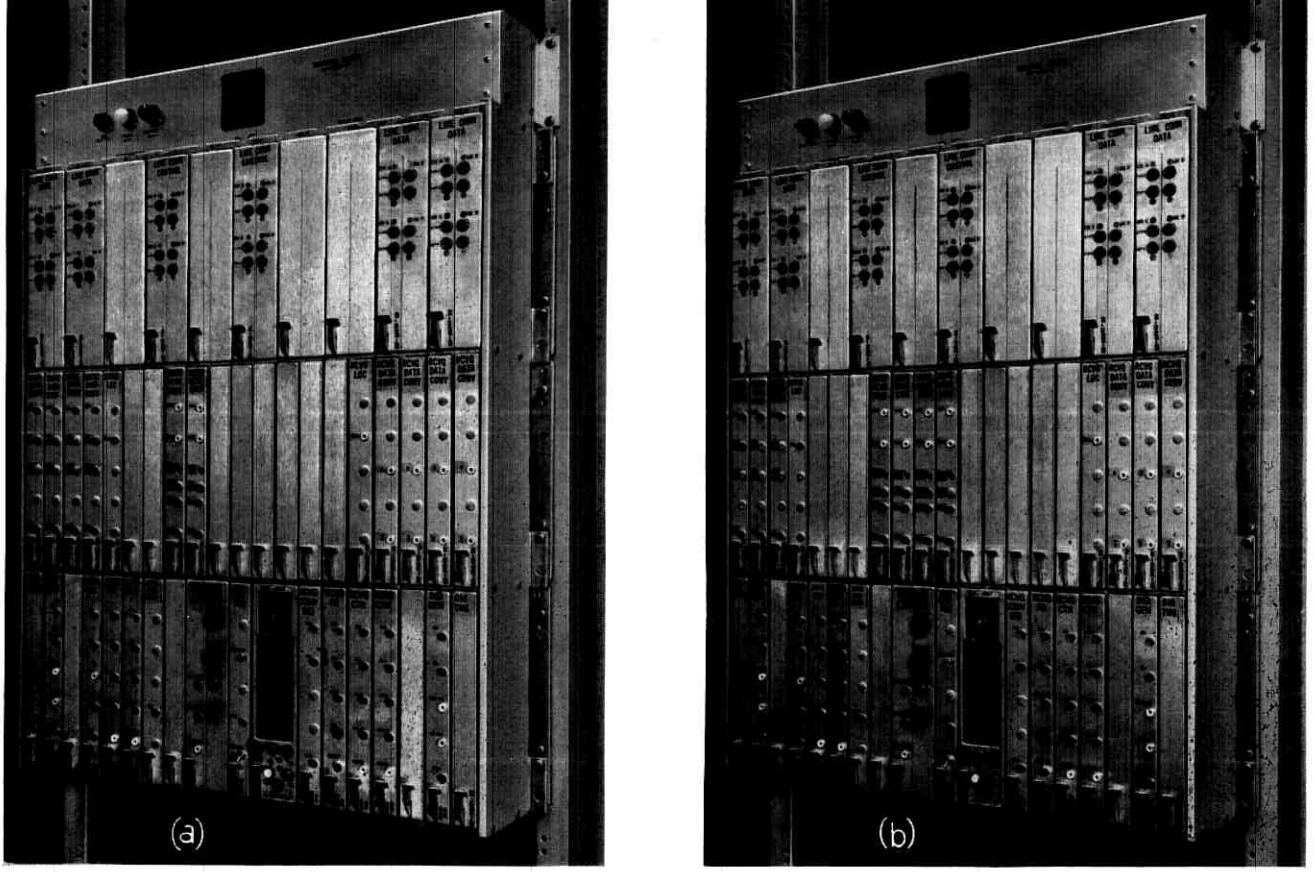


FIG. 18 — Parallel data terminal arranged for (a) 8-bit characters and 2 control channels and (b) 6-bit characters and 4 control channels.

formance is approaching that obtained with a direct tape-to-computer connection.

During this time improvements and changes have been added to the terminal models. Recently both the computer and the terminal equipment were changed to operate at a nominal speed of 90 kilocharacters per second. In addition, changes were made to multiplex the control signals with the data. These control signals were previously transmitted over a separate voice facility.

The data shown on Fig. 19 is indicative of the usage of the network as it has developed over two years. They show the hours per month in computer time that the Murray Hill center received data for processing from the other two locations. The peak, in August, 1964, was the result of an extended down time period of the Holmdel computer for modifications. Recently a second computer and T1 carrier data terminal were added to the Holmdel center. An additional T1 carrier repeatered line facility was provided from Holmdel to Whippany. Terminal-to-line switching equipment allows flexible interconnection of the computers as shown in Fig. 20. It is estimated that for one interconnection alone the Holmdel center is processing at least one-half a computer shift (about 90 hours per month) of work originating at Whippany.

V. GENERAL DEVELOPMENT PROBLEMS

5.1 *Long "0" Sequence Control*

Several general problems arise in applying data signals to T1 carrier lines which must be considered in the design of terminals. One of the

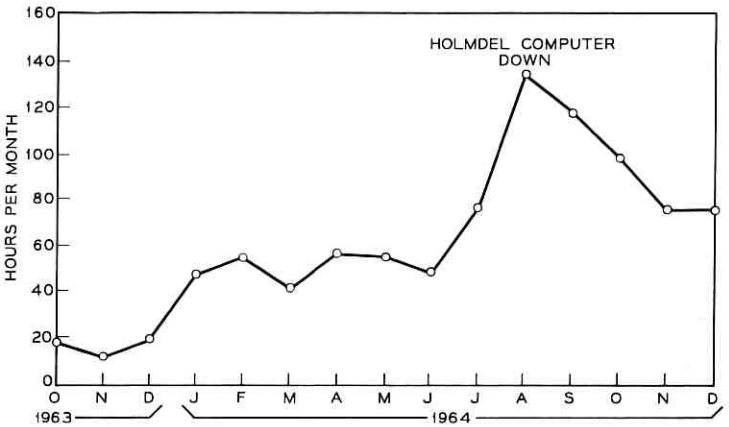


FIG. 19 — Data network usage.

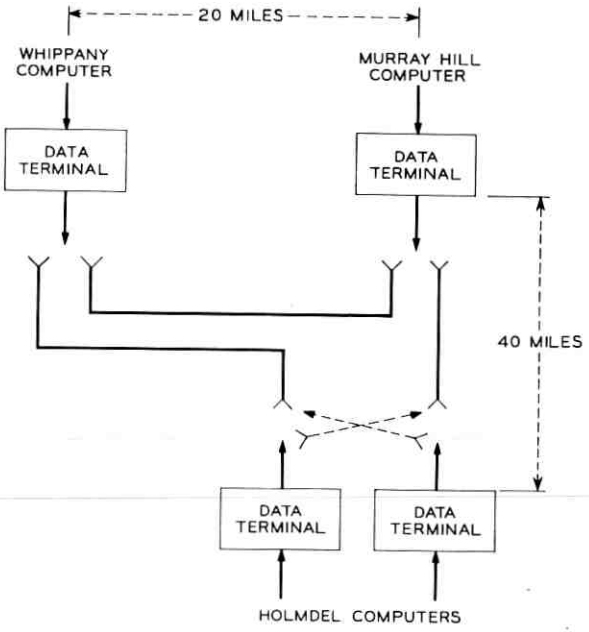


FIG. 20 — Data network layout.

most important relates to limiting long sequences of *ZEROS* in the line signal in order that the repeater clocks remain activated. A properly engineered T1 carrier line should normally tolerate a sequence of up to 14 clock periods without a pulse. In the D1 bank this sequence is limited by the suppression of the all *ZERO* word in the encoder at the loss of only one of the 128 coding levels. A data terminal will not have a similar direct control of the data sequence and other means must be arranged.

In the design of the T1WB-1 bank for asynchronous serial data, use was made of the statistics of the data signal and redundancy in the terminal coding process to control the *ZERO* sequence. Consider for this terminal that the input signals consist of eight statistically independent random binary data signals, synchronous at a 50-kilobit rate. For the worst limiting condition, each data signal has an equal probability of a *ONE* or a *ZERO* in any bit period. It can be shown that, without any constraint on the coding process of the T1WB-1 bank, the probability, P_m, of the occurrence of a continuous sequence of at least m *ZEROS* following a *ONE* is as follows:

m	P_m	*One Occurrence Every*
8	1.23×10^{-5}	0.05 sec.
10	1.90×10^{-6}	0.34 sec.
12	2.92×10^{-7}	2.2 sec.
14	4.50×10^{-8}	14.4 sec.

Although this is several orders of magnitude better than a purely random sequence applied directly on the line, the occurrence of 14 or more *ZEROS* and the likely loss of line synchronization every 15 seconds is not tolerable. A constraint has been added to the coding process of the bank, however, which considerably improves this condition. The Zero Counting Logic Network is provided which counts successive *ZEROS* in the T1 carrier line signal. This is essentially a binary counter which advances on every clock bit but is reset by any *ONE* being transmitted. At a count of 8 *ZEROS* the counter provides a control to all 8 channel logic circuits such that unless a *ONE* is first transmitted for any other reason, the next channel transmitting an "early-late" bit (the 2nd bit of the coding sequence) will transmit a *ONE* whether the transition is *early* or *late*. Forcing the code in this manner produces only a small error in timing but does not produce a complete data bit error. With this added constraint, the following probabilities result:

m	P_m	*One Occurrence Every*
8	1.23×10^{-5}	0.05 sec.
10	2.50×10^{-7}	2.6 sec.
12	5.07×10^{-9}	2.1 min.
14	1.03×10^{-10}	1.8 hours

It is expected that when the traffic statistics are included, a satisfactory line signal condition will result. Any other set of data signal characteristics, such as that resulting from a facsimile signal or the use of an Idle Channel Connector in place of an unequipped channel position, reduces the probability of a long *ZERO* sequence.

Because the statistics as to word format and data rate are more favorable, no *ZERO* sequence control has been incorporated in the design of the terminal for parallel data. This problem can exist only when two successive characters contain an arrangement of 14 or more consecutive *ZEROS* (i.e. 10000000,00000001) and are spaced at the minimum character interval such that no idle line bit (a *ONE* on the line) is transmitted between them. It would be possible, however, to eliminate even this condition by incorporating circuitry which would force an idle line bit between the characters when this condition is anticipated.

These methods of *ZERO* sequence control are dependent on the

particular terminal design and the statistics of the data input signals. Terminal designs for parallel data with large size characters, say 36 bits, or for terminals for hybrid combinations of data and voice signals on one line may require different methods of control. For this reason, consideration is being given to more general arrangements. One method requires that every nth bit, say every 12th bit, on the line be allocated for control purposes. This would result, of course, in a proportionate reduction in data capacity. These bits may be used, however, for auxiliary terminal control functions such as alarm or order wire, or for the transmission of slow speed error control information.

Another method, which may not necessarily reduce the data rate capacity, involves the recoding of the line signal to other than the bipolar format. Although the repeatered line was primarily designed for bipolar signals it has the capability of transmitting ternary signals with certain constraints as to the sequence of pulses of one polarity. Falling somewhat within these constraints is the Paired Selected Ternary (PST) code[5] developed for an experimental high-speed PCM system. Pairs of sequential bits are transposed to pairs of sequential ternary signals following specific rules such that any pair includes at least one positive or negative pulse. Preliminary tests of this code on a T1 carrier repeatered line have shown some degradation of line error rate due to the violation of the bipolar sequence. Further studies are required to determine whether this impairment may be offset by the benefits of the code.

5.2 *Error Rate*

A second general problem concerns T1 carrier system effects on data error rate. The correlation between line error rate and data error rate is dependent upon the data statistics, the terminal coding arrangement, and the distribution of line error rate between errors of omission and errors of commission.

For the series of terminals for asynchronous serial data, previously described, the data error rate may be from $\frac{1}{2}$ to 5 times the line error rate, depending upon the line error distribution. In the case of tape-to-computer data, the computer equipment provides for error checking. If one or more errors occur during transmission of a block of characters, the computer calls for retransmission of the block. In Fig. 21 the probability of a block retransmission, (P_R) is shown as a function of line error rate (P) for several block sizes, (R). It is apparent that line error rates poorer than 10^{-4} will result in little data "throughput" for

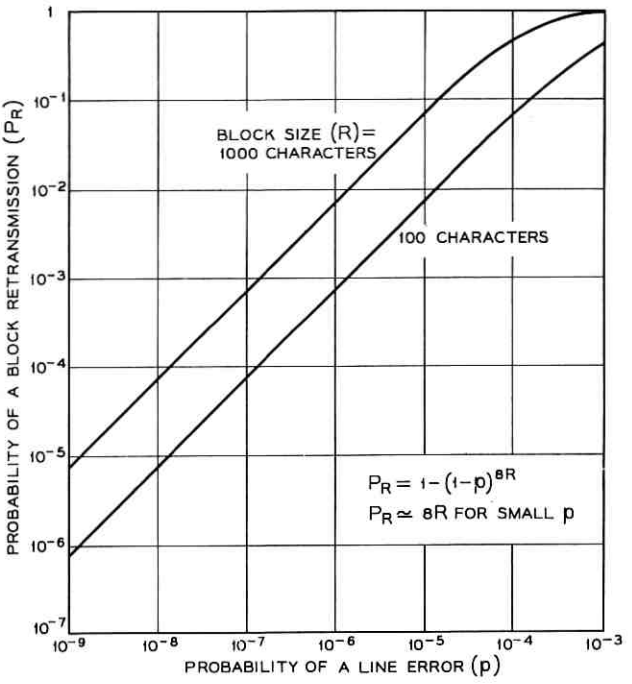


FIG. 21 — Probability of block retransmission.

block sizes of 1000 or more characters. By comparison, adequate, if somewhat degraded, voice frequency performance will result with D1 bank transmission at this error rate.

A study and test program is underway, the objective of which is to provide information on the statistics and distribution of line errors. Although the results are not conclusive, some preliminary tests have indicated that line error rates in the order of 10^{-8} may be expected for properly engineered and installed T1 carrier lines. To attain this capability, however, additional installation testing arrangements are required over those now employed for D1 bank signal transmission. For example, cases have occurred in which repeaters have been installed with incorrect line build-out networks, an error which was not discovered in the final testing procedures or by the terminal tests of the D1 bank. Under these conditions the line is extremely sensitive to line signal patterns which typically may be generated by the parallel data terminal. A simple test set generating such a pattern may provide one important test facility.

5.3 *Line Pattern Controls*

Associated with the error rate problem is the problem of data terminals producing certain fixed patterns on the line which may interfere with the timing of other systems. These patterns may be due to the data signal or due to the terminal coding process. This is of importance only in large cross-section systems on a single cable where lines carrying these interfering patterns make up a large part of this cross-section. Some control of pattern density may be obtained by selective code inversion of certain of the data signal bits in the transmitting terminal under the control of terminal timing. These would then be re-inverted in the receiving terminal. Such a change must be correlated, of course, with the control of the all-*ZERO* sequence.

VI. POTENTIAL TERMINAL APPLICATIONS

6.1 *Hybrid Terminal Arrangements*

The arrangements for multiplexing several like data signals has been described for the T1WB-1 and -2 wideband banks for serial data. These arrangements may be easily extended to include the multiplexing of different types of data signals which are processed into similar formats on the T1 carrier line. For example, a parallel data signal of the tape-to-computer type requiring only one-half the line capacity may be converted into alternate bits on the line. The remaining bit capacity may then be used for a combination of serial data signals requiring only one-half the line capacity. The timing circuits for this hybrid terminal arrangement may be common to both sets of equipment or separate timing circuits may be used if they are synchronized. It is clear, however, that data signals in this format on the line do not coordinate with the 8-bit word format of the D1 bank. Consider however, the conversion of the D1 bank signals into the format of the data signals. In Fig. 22, a D1 bank is shown, equipped with 12 channels in the even channel group. The output is transferred into a register which, under the control of timing, distributes the bits of the signal into alternate *EVEN* bits in the T1 carrier line. By synchronizing and framing the data timing circuits with this signal, the 12 voice channels may be interleaved with data signals occupying the *ODD* line bits. At the receiving end the *EVEN* bits are selected and transferred into a similar register where they are rearranged into the 8-bit word format.

An alternate arrangement for multiplexing voice and data signals may be obtained by converting the data signals into the D1 bank for-

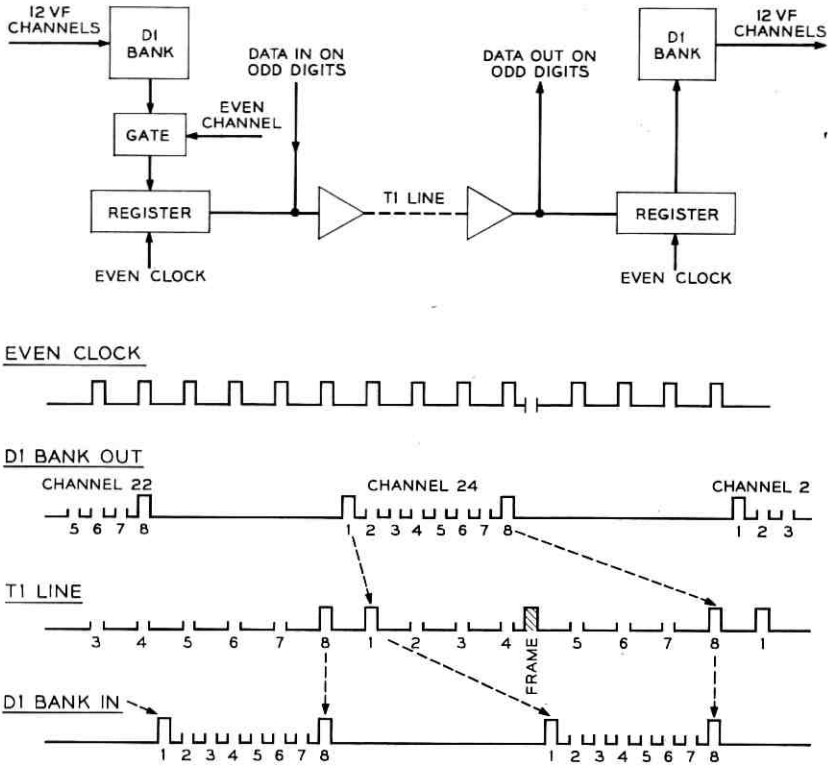


FIG. 22 — Multiplexing with a D1 Bank.

mat using similar register circuits in the reverse order to that described. Although the data would then appear in an 8-bit word group, and hence could be multiplexed with 12 channels from the D1 bank, it is not constrained in this format. That is, the organization of information in the data word bears no direct relationship with the channel word.

The advantage of one of these arrangements over the other is somewhat dependent upon the potential application. For the case where, say, three voice signals are to be replaced with one 50-kilobit serial data signal, it can be seen that less register storage capacity is required if the second alternative is applied. In this alternative the data signal is converted rather than the voice signals from the D1 bank. It is unlikely that a few voice signals from the D1 bank would be multiplexed with a majority of data signals. The cost of the D1 bank is such as to make it more attractive to transmit these few voice signals as part of a large group of voice signals over another T1 carrier line.

The first arrangement in which the data remains distributed bit-by-bit on the line, would appear to have one advantage, however. From the preliminary tests on line error distribution there is some evidence that a correlation in occurrence of errors exists such that there is a likelihood of errors occurring in adjacent or near adjacent time slots. By leaving multiplexed data in the distributed format, there is less likelihood of the event of paired errors occurring in only one channel, improving the effectiveness of simple error control coding which may exist in that channel signal.

6.2 *Synchronous Serial Data Terminals*

The terminals previously described for asynchronous serial data are, of course, capable of handling synchronous serial data at any rate up to a maximum rate determined by timing options. The penalty paid for this high degree of flexibility, however, is efficiency since three T1 carrier line bits are required for each data bit. The present conditions in the data field as to the variety of data transmission requirements and the economics of T1 carrier relative to other facilities in the short-haul plant justify the use of this approach. However, if standard fixed rates for synchronous data transmission develop, terminals designed specifically for these fixed rates may be desirable. It is clear that for specific synchronous rates, efficiencies approaching one data bit per line bit are feasible.

The two approaches which are considered here do not require synchronization and phasing of the data source with the T1 carrier terminal clock. They do require, however, that the data be at specific rates within certain tolerance limits.

In the discussion of transmission of the parallel data terminal it was shown that large parallel characters can be put onto the T1 carrier line with fairly high efficiencies from the standpoint of T1 carrier bits required per data bit. This is due to the fact that only one timing or index bit is required to provide accurate timing for all of the bits of the character. This approach may be extended to synchronous serial data. Consider a synchronous serial data signal which is shifted in alternate groups of say 8 bits into 2 storage registers. These groups are transferred alternately into the parallel data terminal which transmits them on the T1 carrier line at a bit rate slightly higher than the original data bit rate. The average group rate on the line would be the same as the data group rate or $\frac{1}{8}$ the data bit rate. At the receiving end the data groups are transferred to a set of registers. The serial data is then read out of these registers under control of a "smoothing" clock

which derives its frequency control from the marking of the group rate or from a servo loop controlled by the status of storage in the registers.

One of the most promising approaches to efficient transmission of synchronous serial data signals involves the application of a technique proposed by J. S. Mayo for network synchronization of high-speed PCM systems.[6] Briefly, a synchronous serial bit stream of rate f_B may be processed into a line or fraction of a line with a capability f_L, slightly higher than f_B. By means of an elastic store and logic circuits, additional bits are "stuffed" into the bit stream f_B as necessary to make its rate identically f_L. Additional information is included on the line to identify the "stuffed" bits. At the receiving end the stuffed bits are removed, and the signal is smoothed in an elastic store. This technique will allow efficiencies approaching 100 per cent.

6.3 *T1 Carrier and High-Speed Digital Transmission Systems*

A plurality of wideband data signals, such as the 50-kilobit signal discussed earlier, can be transmitted over a T1 carrier repeatered line. In addition, the line has the capacity for considerably higher data rates, up to 1.5 megabits for synchronous serial signals.

T1 carrier was designed for application in the short haul plant. Presently, the signals transmitted via a T1 carrier end link must be brought down to baseband and transformed by analog terminals for long haul transmission. One may expect that digital transmission techniques will be adopted in the toll plant in the future, thus providing a broader use for a variety of digital data terminals such as those considered in this article.

VII. ACKNOWLEDGMENTS

The work described in this article is the result of the efforts of many in the engineering and development organizations of Bell Laboratories. The writers particularly wish to acknowledge the contributions of Messrs. R. G. DeWitt and J. P. Forde, whose planning and evaluation of the experimental service among the Bell Laboratories computation centers provided important data for the development program.

REFERENCES

1. Fultz, K. E., and Penick, D. B., The T1 Carrier System, B.S.T.J., *44*, Sept., 1965, pp. 1405–1451.
2. Gravis, H., and Crater, T. V., Engineering of T1 Carrier System Repeatered Lines, B.S.T.J., *42*, March, 1963, pp. 431–486.
3. Davis, C. G., and Thomas, L. C., Patent No. 1394485 (France), February 22, 1965.
4. Becker, F. K., Davey, J. R., and Saltzberg, B. R., Conference Paper, A.I.E.E. Fall General Meeting, Detroit, Michigan, October, 1961.
5. *Transmission Systems for Communications*, Chapter 26.
6. Mayo, J. S., PCM Synchronization, U. S. Patent No. 3,136,861, June 9, 1964.

An Algorithm for Solving Nonlinear Resistor Networks

By JACOB KATZENELSON

(Manuscript received May 17, 1965)

This article describes an algorithm for solving electrical networks which consist of linear and nonlinear resistors and independent sources, and where the characteristics of each of the resistors is described by a function $G_k(\cdot)$, $i = G_k(v)$, *where* $G_k(\cdot)$ *is continuous, monotonically increasing, piecewise linear, and one-to-one from* $(-\infty, \infty)$ *onto* $(-\infty, \infty)$, *and where* k *is an index which spans all the resistors in the network.*

The solution is found by solving successively equivalent linear networks which represent the nonlinear network locally and which correspond to a "solution curve." Essential to the efficiency of the computation process is the method of modifying matrices which enables the process to find the inverse of a conductance matrix by modifying another matrix rather than by matrix inversion.

The algorithm provides a fast computation method for both of the following two cases: (1.) the network contains both linear and nonlinear resistors and (2.) the sources are functions of time and the solution is required for successive values of time. In the latter case the algorithm computes each solution from the previous one rather than solving each case independently.

I. INTRODUCTION

This article considers an algorithm for solving electrical networks which consist of linear and nonlinear resistors and independent sources and where the current-voltage relation of each of the nonlinear resistors is described by a function $G_k(\cdot)$, $i = G_k(v)$, where $G_k(\cdot)$ is continuous, monotonically increasing, piecewise linear, and one-to-one from $(-\infty, \infty)$ onto $(-\infty, \infty)$, and where k is an index which spans all the resistors in the networks.

A piecewise linear network of this type can be considered to be an approximation to a more general nonlinear resistor network where the corresponding function $G_k(\cdot)$ is continuous, monotonically increasing, and one-to-one from $(\infty, -\infty)$ onto $(-\infty, \infty)$ but not necessarily piecewise linear.

Networks of the last type were discussed by various authors,[1,2,3,4,6,7] in particular Duffin, who has shown[1] that such networks have a solution which is unique. Various methods were proposed for finding the numerical value of the solution. Birkhoff and Diaz[2] gave an iterative method similar to Seidel's method[5] which is a form of relaxation procedure for solving linear equations. A direct iterative method[5] similar to the "standard" method of solving linear equations by iteration was described by Katzenelson and Seitelman.[6] An exact method (convergence in a finite number of steps) was described by Minty.[7] This latter method approximates a monotonic increasing characteristic by "stairs" and solves the approximating network by a search procedure.

The algorithm described here approximates the nonlinear resistors by piecewise linear resistors. The solution is found in a finite number of steps by successively solving linear networks, which locally represent the original network, along a path which is called the "solution curve" (Section III). Essential to the efficiency of the computation process is (6) by which the inverse of a conductance matrix is found by modifying another matrix rather than by matrix inversion.

In comparison with other solution methods, [2,6,7] the advantage of the algorithm is in providing a fast computation method for the cases where (a) the network contains both linear and nonlinear resistors and where (b) the sources are time dependent and the solution is required for all time t in some interval $[t_0, t_\alpha]$. The iteration methods, [2,6] and Minty's method,[7] do not take direct advantage of the occurrence of linear resistors in the network. In this algorithm, however, these resistors simplify the computation considerably. Case (b) is solved by sampling the time interval and for each instant of time solving the networks with constant sources whose value is equal to the value of the time-varying sources at that instant. The algorithm computes each solution from the previous one in a rather simple manner resulting in a significant reduction of computation time. Our interest in a fast calculation method for resistive networks with time dependent sources is related to the problem of solving nonlinear *RLC* networks. It was shown[4] that these networks can be viewed as a combination of three one-element-kind networks: a capacitive network, a resistive network and an inductive network. Generally, finding the time response of the *RLC* network involves solving, for each instant of time t, the three one-element-kind networks. Each of these networks is solved as a purely resistive network with the currents or voltages of the other networks playing the role of sources. An algorithm of the type described here can be used to obtain efficiently the solution of each one-element-kind network at time t from its solution at $t - \Delta t$.

Properties of piecewise linear network which are relevent to the algorithm are discussed in Section II. Section III contains a general description of the algorithm. Section IV contains a convergence proof. Section V considers the computation time which is required for solving nonlinear resistor networks. The various methods of solution are compared from this point of view. Appendix A describes a modification which reduces the computation time used by the algorithm. A step-by-step description of the algorithm is given in Appendix B.

II. PROPERTIES OF PIECEWISE LINEAR NETWORKS

This section describes the type of networks which are solved by the algorithm. A network η can be solved by the algorithm if it satisfies the following conditions:

(1.) η consists of a finite number of branches. Each branch is either a resistor (linear or nonlinear) or an independent source. Without loss of generality it can be assumed that η is connected, is nonseparable and does not contain cut sets of current sources only or loops of voltage sources only. It follows from the above that η contains a tree τ such that all voltage sources are tree branches and all current sources are links. Without loss of generality it can be further assumed that each current source appears in parallel with a resistive tree branch, and that each resistive tree branch has only one current source connected in parallel with it.[4]

(2.) The characteristics of each of the resistors of η satisfy the following conditions:

(a.) Let i and v be the current and the voltage of the resistor. The sign convention of i and v is shown in Fig. 1. The characteristics of the resistor can be represented by a function $G(\cdot)$, $i = G(v)$, where $G(\cdot)$ is a continuous, *piecewise linear*, monotonic increasing function which is one-to-one from $(-\infty, \infty)$ onto $(-\infty, \infty)$. (Fig. 2).

(b.) Each characteristic has a finite number of breakpoints in any finite interval.

It follows from (a.) and (b.) that in each finite interval the slope of $G(\cdot)$ is bounded from above and below. The lower bound is positive.

Let us discuss the properties of networks which satisfy (1.) and (2.). It follows from Duffin's work[1] that networks of this type have a unique solution. Thus, to any value of the sources corresponds one and only

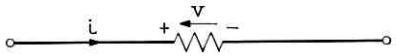


Fig. 1 — Sign convention for a branch of the network.

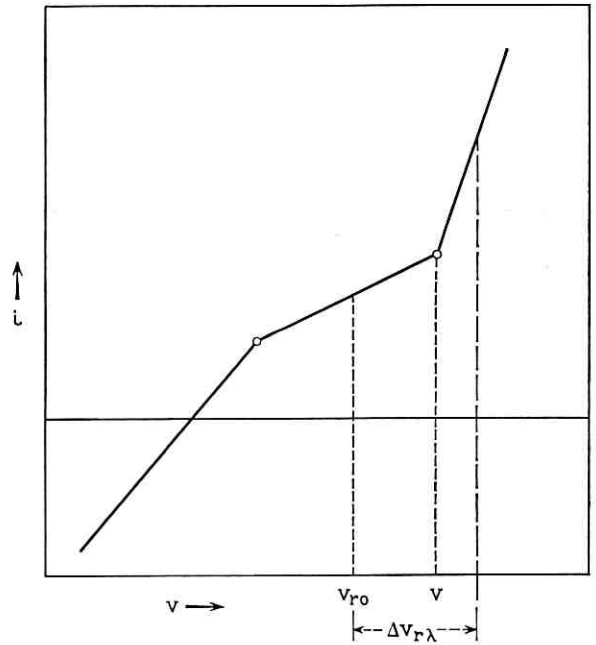


Fig. 2 — A monotonically increasing piecewise linear resistor characteristic.

one set of voltages and currents of the resistors which satisfies the Kirchhoff's laws and the branch characteristics. The purpose of our algorithm is to evaluate these voltages and currents for a given value of the sources. Before proceeding let us make a few notations. We shall choose a tree of the network and refer to branches which are in the tree as tree branches and branches which are not in the tree as links. Let τ be a tree of η such that all voltage sources are tree branches and all current sources are links of τ. Let $\mathbf{v}_r(\mathbf{i}_r)$ be a vector whose components are the voltages (the currents) of the resistive branches. Let $\mathbf{e}_r$ denote the vector whose components are the voltages of the resistive tree branches. Similarly, $\mathbf{E}$ and $\mathbf{J}$ denote the voltage and current sources of the network.

From the fact that η has a unique solution for any value of $(\mathbf{E},\mathbf{J})$, it follows that there exists a (single valued) function which maps $(\mathbf{E},\mathbf{J})$ into $\mathbf{e}_r$. The domain of this function is $\mathcal{E} \times \mathcal{J}$ where $\mathcal{E}$ and $\mathcal{J}$ denote the vector spaces corresponding to the domains of $\mathbf{E}$ and $\mathbf{J}$, respectively.

Similarly, it follows from conditions (1.) and (2.) that $\mathbf{E}$ and $\mathbf{e}_r$ determine uniquely the currents in all the resistors and since the com-

ponents of $\mathbf{J}$ are the fundamental cut set current sources it follows that there exists a (single valued) function which maps $(\mathbf{E},\mathbf{e}_r)$ into $\mathbf{J}$. The domain of this function is $\mathcal{E}_r \times \mathcal{E}$ where $\mathcal{E}_r$ is the vector space corresponding to the domain of $\mathbf{e}_r$.

Let us fix $\mathbf{E}$. It follows from the above that the mapping $\mathbf{J} \rightarrow \mathbf{e}_r$, for a given $\mathbf{E}$, is one-to-one from $\mathcal{J}$ onto $\mathcal{E}_r$. In addition, it was proved[4] that for networks satisfying conditions (1.) and (2.) the mappings $\mathbf{J} \rightarrow \mathbf{e}_r$ and $\mathbf{e}_r \rightarrow \mathbf{J}$ are continuous and satisfy Lipschitz conditions on any bounded set of their respective domains.

The network η of Fig. 3 will be used for illustrating the next property of interest. Let η satisfy conditions (1.) and (2.). The tree τ of η consist of three branches 1, 2, and 3. Since 3 is a voltage source $\mathbf{e}_r$ and $\mathbf{J}$ are two dimensional vectors and the spaces $\mathcal{E}_r$ and $\mathcal{J}$ are planes.

Denote the voltage corresponding to the kth breakpoint of the first resistor by e_1^k. Consider the space $\mathcal{E}_r$ and the locus of all points $\mathbf{e}_r$ such that the voltage on the first resistor, e_{r1}, is equal to the voltage corresponding to one of the breakpoints of the resistor characteristics, e_1^k, $k = 1, 2, \cdots, n$. This locus is a set of straight lines parallel to the e_{r2} axes. Similarly the locus corresponding to the breakpoints of the second resistor are lines parallel to the e_{r1} axes. The lines corresponding to the third resistor satisfy $E + e_{r1} - e_{r2} = e_4^k$ which are lines which form a 45° angle with the axes. These lines partition the plane $\mathcal{E}_r$ into regions as shown in Fig. 4. In the case of larger networks these loci are suitable hyperplanes which partition the space $\mathcal{E}_r$ into similar regions.

The interesting fact about the regions is that inside each, the mapping $\mathbf{e}_r \rightarrow \mathbf{J}$ is linear. This property is heavily used by the algorithm.

With each region we associate a linear network called the linear

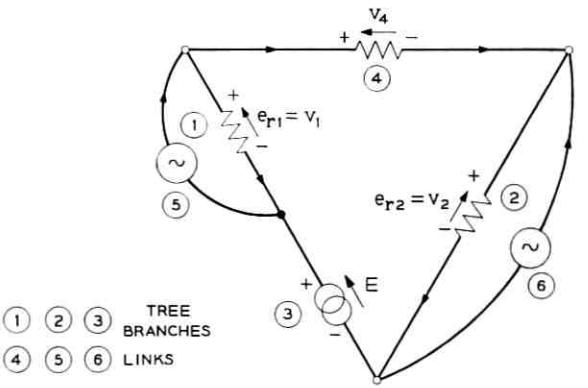


Fig. 3 — Resistive network.

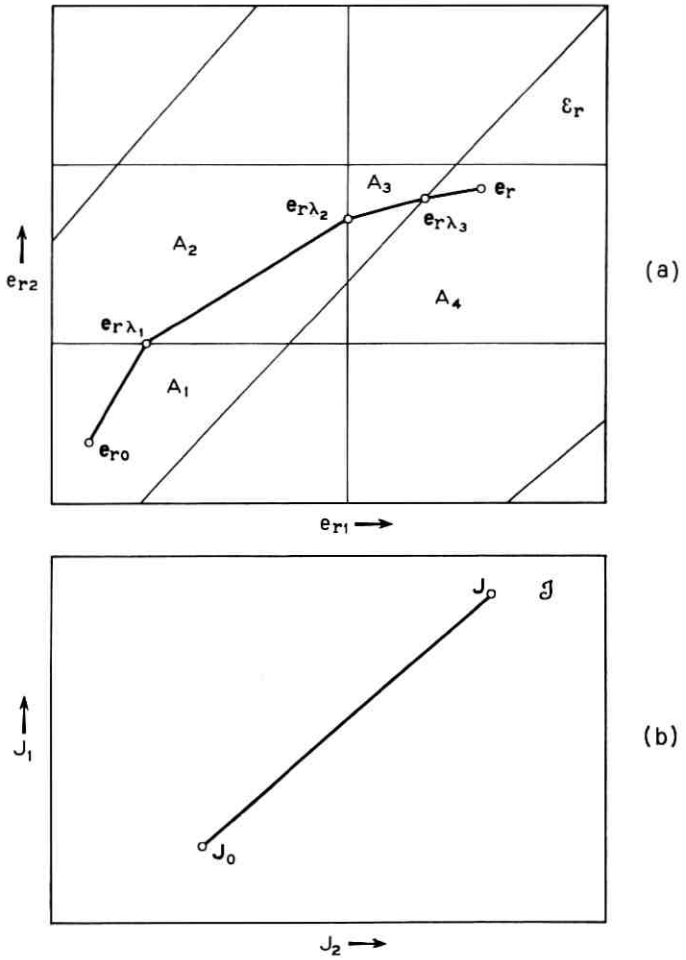


Fig. 4 — The partition of $\mathcal{E}_r$ into linear regions and the solution curve.

equivalent network of the region. It has the same topology as the original network with each nonlinear resistor replaced by a linear resistor with conductance equal to the incremental conductance of the nonlinear resistance at the region.

Before proceeding with the algorithm let us describe a convenient notation. Consider the network of Fig. 3. The cut set equation could be written as

$$g_1(e_{r1}) + g_4(e_{r1} + E - e_{r2}) = J_5$$
$$g_2(e_{r2}) - g_4(e_{r1} + E - e_{r2}) = J_6$$

where g_1, g_2, and g_4 are the functions which correspond to the characteristics of the resistors. This can be written symbolically as

$$\mathbf{G}^* \begin{pmatrix} \mathbf{e}_r \\ \mathbf{E} \end{pmatrix} = \mathbf{J};$$

where $\mathbf{G}$ is a mapping from $\mathcal{E} \times \mathcal{E}_r$ to $\mathcal{J}$ and $\begin{pmatrix} \mathbf{e}_r \\ \mathbf{E} \end{pmatrix}$ denotes a vector whose components are the components of $\mathbf{e}_r$ and $\mathbf{E}$ arranged as the notation implies, namely components of $\mathbf{e}_r$ first and components of $\mathbf{E}$ second. This notation is used as a shorthand notation for the network equations. (Somewhat more elaborate notation for equations of this type is discussed in Ref. 4).

This concludes the discussion of the properties of piecewise linear resistor networks. The algorithm itself will be discussed next.

III. A GENERAL DESCRIPTION OF THE ALGORITHM

The following contains a general description of the algorithm. A detailed, step-by-step description will appear in Appendix B.

The solution of the network problem requires the evaluation of $\mathbf{e}_r$ from

$$\mathbf{G}^* \begin{pmatrix} \mathbf{e}_r \\ \mathbf{E} \end{pmatrix} = \mathbf{J}. \tag{1}$$

Consider the spaces $\mathcal{E}_r$ and $\mathcal{J}$. Let us chose a point in $\mathcal{E}_r$ and denote it by $\mathbf{e}_{r0}$. The corresponding point in $\mathcal{J}$ is given by

$$\mathbf{G}^* \begin{pmatrix} \mathbf{e}_{r0} \\ \mathbf{E} \end{pmatrix} = \mathbf{J}_0. \tag{2}$$

Consider the points $\mathbf{e}_{r\lambda}$ which are a solution to

$$\mathbf{G}^* \begin{pmatrix} \mathbf{e}_{r\lambda} \\ \mathbf{E} \end{pmatrix} = \mathbf{J}_0 + \lambda(\mathbf{J} - \mathbf{J}_0) \tag{3}$$

where λ attains all values between 0 and 1.

For $0 \leqq \lambda \leqq 1$, the right hand side of (3) describes a straight line in $\mathcal{J}$ which connects the point $\mathbf{J}_0$ with $\mathbf{J}$. For a given $\mathbf{E}$, (3) describes a mapping of this line from $\mathcal{J}$ to $\mathcal{E}_r$. The image of the line $(\mathbf{J}_0, \mathbf{J})$ in $\mathcal{E}_r$ space has the following properties: for $\lambda = 0$, $\mathbf{e}_{r\lambda} = \mathbf{e}_{r0}$; for $\lambda = 1$, $\mathbf{e}_{r\lambda} = \mathbf{e}_r$; which is the solution of (1). As the mapping is continuous and one-to-one onto *the line* will be mapped to a path from $\mathbf{e}_{r0}$ to $\mathbf{e}_r$. As within each region the mapping from $\mathcal{J}$ to $\mathcal{E}_r$ is linear, inside each region of $\mathcal{E}_r$ the path consists of a linear segment. An example of such a path is given in Fig. 4. This path is called *the solution curve.*

Generally speaking the algorithm calculates the solution of the nonlinear network by tracing the solution curve from its beginning at the point chosen arbitrarily $\mathbf{e}_{r\lambda} = \mathbf{e}_{r0}$, $\lambda = 0$, to its end, $\mathbf{e}_{r\lambda} = \mathbf{e}_r$, $\lambda = 1$ which is the solution of (1). This operation will be explained in the following.

The solution curve is traced by taking advantage of the fact that inside each region the mapping is linear. Consider the example of Fig. 3 and its corresponding $\mathcal{E}_r$ and $\mathcal{J}$ spaces of Fig. 4. In Fig. 4(a), the regions A_1, A_2, etc., are the regions through which the solution curve passes and the points $\mathbf{e}_{r\lambda_1}$, $\mathbf{e}_{r\lambda_2}$ etc., are the intersections of the solution curve with the boundaries. The point $\mathbf{e}_{r\lambda_1}$ is on the boundary of region A_1 which includes the initial point $\mathbf{e}_{r0}$. The point $\mathbf{e}_{r\lambda_1}$ can be calculated from $\mathbf{e}_{r0}$ as follows. Let λ_1 be the value of λ corresponding to $\mathbf{e}_{r\lambda_1}$. Let $\mathbf{G}_{A_1}$ denote the conductance matrix of the linear equivalent network for region A_1.

$$\mathbf{e}_{r\lambda_1} - \mathbf{e}_{r0} = \lambda_1 \mathbf{G}_1^{-1}(\mathbf{J} - \mathbf{J}_0). \tag{4}$$

To find $\mathbf{e}_{r\lambda_1}$, Δe_r is first evaluated from

$$\Delta \mathbf{e}_r = \mathbf{G}_1^{-1}(\mathbf{J} - \mathbf{J}_0). \tag{5}$$

Next, we find the largest M, $0 \leqq M \leqq 1$ such that $\mathbf{e}_{r0} + M\Delta\mathbf{e}_r$ is in region A_1.
Thus,

$$\mathbf{e}_{r\lambda_1} = \mathbf{e}_{r0} + M\Delta\mathbf{e}_{r0}.$$

The actual computation of M is described in detail in Appendix B, steps 5 and 6. Now $\mathbf{e}_{r\lambda_1}$ is considered to be a part of A_2 and $\mathbf{e}_{r\lambda_2}$ is calculated from $\mathbf{e}_{r\lambda_1}$ in the same way as $\mathbf{e}_{r\lambda_1}$ was calculated from $\mathbf{e}_{r0}$. The process continues in this way and proceeds along the solution curve until $M = 1$ indicates that $\lambda = 1$ and that the solution is found.

At this point let us consider the computational aspects of the algorithm. The network of Fig. 3 and the corresponding Fig. 4 will be used again for illustration.

At each region, (5) is used to find the intersection of the solution curve with the region boundary. The use of (5) involves the inverse of the conductance matrix of the corresponding region. In our example, the $\mathbf{e}_{r\lambda_1}$ is obtained from $\mathbf{e}_{r\lambda_0}$ and $\mathbf{e}_{r\lambda_2}$ from $\mathbf{e}_{r\lambda_1}$ by using the matrices $\mathbf{G}_{A_1}^{-1}$ and $\mathbf{G}_{A_2}^{-1}$ which correspond to the conductance matrices of the equivalent linear network in regions A_1 and A_2. The main point is that the process of obtaining the conductance matrix and inverting it directly is slow in comparison with all other operations in the algorithm and

the required time increases rapidly with the size of the network. The algorithm circumvents this difficulty as follows: The inverse of the matrix of the new region is obtained by modifying the inverse of the matrix of the previous region. Consider again our example:

The algorithm computes $\mathbf{G}_{A_1}^{-1}$ in the first step. However, once a boundary is crossed $\mathbf{G}_{A_2}^{-1}$ is obtained from $\mathbf{G}_{A_1}^{-1}$ by the method of modifying matrices.[8,9,10] This is a method of inverting a matrix which is a modification of another matrix whose inverse is known. The formula involved is

$$(\mathbf{F} + \mathbf{IHK})^{-1} = \mathbf{F}^{-1} - \mathbf{F}^{-1}\mathbf{I}(\mathbf{KF}^{-1}\mathbf{I} + \mathbf{H}^{-1})^{-1}\mathbf{KF}^{-1} \tag{6}$$

where $\mathbf{F}$, $\mathbf{I}$, $\mathbf{H}$ and $\mathbf{K}$ are matrices of suitable dimensions and the inverse of $\mathbf{H}$ is assumed to exist. Note that if $\mathbf{I}$ is a column vector, $\mathbf{K}$ a row vector, $\mathbf{H}$ a 1×1 matrix and $\mathbf{F}^{-1}$ known then the calculation of the left hand side requires the inversion of a 1×1 matrix only. The application of (6) to *linear* networks is quite well known.[8] It is related to the Kron method and is used for finding the inverse of conductance matrices, adding resistors and nodes to a network, etc. The application of (6) to nonlinear networks is believed to be new.

In the example of Fig. 3, assume that while moving from A_1 to A_2 only one boundary was crossed which implies that the two equivalent linear networks differ in the conductance of one resistor only. Let this be the second resistor and let this difference be ΔG. Thus,

$$\mathbf{G}_1 = \mathbf{Q}\begin{bmatrix} G_1 & 0 & 0 \\ 0 & G_2 & 0 \\ 0 & 0 & G_3 \end{bmatrix}\mathbf{Q}_T$$

where $\mathbf{Q}$ is the fundamental cut set matrix, after voltage sources are replaced by short-circuit and current sources removed from the network. The subscript T denotes transposition. Then

$$\mathbf{G}_2 = \mathbf{Q}\begin{bmatrix} G_1 & 0 & 0 \\ 0 & G_2 + \Delta G & 0 \\ 0 & 0 & G_3 \end{bmatrix}\mathbf{Q}_T = \mathbf{G}_1 + \mathbf{Q}\begin{bmatrix} 0 & 0 & 0 \\ 0 & \Delta G & 0 \\ 0 & 0 & 0 \end{bmatrix}\mathbf{Q}_T$$

$$= \mathbf{G}_1 + \mathbf{Q}_{c2} \cdot \Delta\mathbf{G} \cdot [\mathbf{Q}_{c2}]_T$$

where $\mathbf{Q}_{c2}$ is a column vector equal to the second column of Q. Thus to find $\mathbf{G}_{A_2}^{-1}$ from $\mathbf{G}_{A_1}^{-1}$, (6) can be used with

$$\mathbf{I} = \mathbf{Q}_{c2}, \qquad \mathbf{K} = [\mathbf{Q}_{c2}]_T, \qquad \mathbf{F}^{-1} = \mathbf{G}_{A_1}^{-1} \quad \text{and} \quad \mathbf{H} = \Delta\mathbf{G}.$$

It is to be noted that when one boundary is crossed, the use of (6) requires the inversion of a 1×1 matrix only since both $\mathbf{H}$ and $\mathbf{KF}^{-1}\mathbf{I}$ are 1×1.

Thus, a difficult matrix inversion is performed only once for the initial region in which the computation starts. Inverses of matrices corresponding to other regions along the solution curve are computed successively by modifying the matrix corresponding to the previous region. This process takes only a small fraction of the time required for a direct inversion and essentially makes the algorithm as described above a usable computation process.

When the solution curve intersects a boundary it always continues to the adjacent region. The identity of the adjacent region is quite clear in Fig. 4 where the solution curve crosses one boundary at a time. Fig. 5 illustrates the case where the solution curve passes through the intersection of two boundaries and there are three adjacent regions. In order to apply (6) to this case it is necessary to find the region in which the solution curve continues behind the double boundary point. The region can be found by a search procedure which selects a region and attempts to continue the curve in it. If the attempt fails the next region is selected.

The occurrence of a solution curve intersecting a multi-boundary point is admittedly rare. However, when it occurs the search procedure can be quite lengthy for large networks. Large networks can have points in which a large number of boundaries intersect forming a large number of adjacent regions namely $2^n - 1$ where n is the number of boundaries which intersect in one point. Appendix A describes a method to overcome this difficulty.

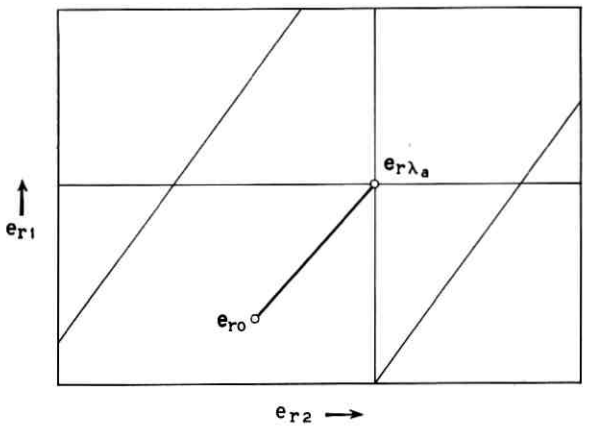


Fig. 5 — Crossing of a boundry intersection.

IV. CONVERGENCE OF THE ALGORITHM

In this section, it is proved that the algorithm converges in a finite number of steps. It will be proved that the solution curve (Fig. 4) crosses a finite number of boundaries and therefore this algorithm converges in a finite number of steps.

For simplicity, consider the network of Fig. 3. In $\mathcal{J}$ space the image of the solution curve is a straight line joining $\mathbf{J}_0$ and $\mathbf{J}$. Consider the boundaries of a region, say A_1, (Fig. 4(a)) in $\mathcal{E}_r$. These boundaries are linear segments and since the mapping $\mathcal{E}_r \rightarrow \mathcal{J}$ is continuous and linear inside each region the image of each segment is a linear segment. Now, if a segment and the solution curve have some points in common, one of two possibilities exists: (1.) they have one common point or (2.) they have a common finite interval. The first case corresponds to one boundary crossing. When a segment of this type is crossed the solution line never crosses it again. In the second case, the solution curve remains in one region since each region includes its boundaries.

Let $\| \mathbf{A} \|$ denote the square norm of $\mathbf{A}$ and for a given $\mathbf{J}$, $\mathbf{J}_0$ let S_1 be the set

$$\{\mathbf{J}_1 \mid \| \mathbf{J}_1 - \mathbf{J}_0 \| < \| \mathbf{J} - \mathbf{J}_0 \| + \varepsilon, \qquad \varepsilon > 0\}.$$

Let S_2 be the image of S_1 in $\mathcal{E}_r$. S_2 is bounded since S_1 is bounded and the mapping is continuous. It follows from the last part of condition (2.), Section II, that S_2 contains a finite number of regions. Therefore, S_1 contains a finite number of segments. Now the image in $\mathcal{J}_r$ of the solution curve is the linear segment $(\mathbf{J},\mathbf{J}_0)$. As it is linear and included in S_1 it can cross each boundary segment only once; it therefore has a finite number of crossing points. Therefore the solution terminates in a finite number of steps.

V. COMPUTATION TIME

This section is concerned with the computation time required by the algorithm and with the way this time is used by various parts of the algorithm. The times quoted corresponds to a FORTRAN program written by the author and run on the IBM 7094.

From the nature of the algorithm, it is clear that the computer time is a function of the size of the network and the distance of the initial point from the solution. The size of a network, whose resistors are all nonlinear, was defined as the number of nonlinear resistors. As the distance between initial point and result changes from problem to problem, the following parameters rather than the total solution time were used to investigate the dependency of the computation time on the size of

the network: (i) set up time — most of it is the time to invert the conductance matrix, (ii) the time used by the solution curve to cross a region from one boundary to the other, and (iii) the time used by the solution curve to reach from a boundary of a region which contains the solution to the solution itself. While (i), (ii), and (iii) depend on the size of the network, they do not depend on the initial point or the particular values of the resistors. The results for (ii) are given in Table I. Parameters (ii) and (iii) are approximately equal, and if (6) is used for setting and inverting the (i) is approximately equal to (ii) times the number of resistors in the network.

VI. CONCLUSION

We shall conclude the article by considering again the properties of the algorithm and by comparing the algorithm with the iteration method of Ref. 6.

Let us again consider the solution of the nonlinear resistor network with time-varying sources where the solution is required on some time interval $[t_0, \alpha t]$. Once the solution is found for t_0, the solution for $t_0 + \Delta t$ can be found by modifying the conductance matrix which has been used to get the solution for t_0. Thus, the solution at each instant of time is obtained by modifying the results of previous calculation rather than setting up an independent calculation and a significant reduction of computation time is obtained.

Another problem in which the algorithm can be used advantageously is finding a solution of a network which was obtained from another network by adding or subtracting branches. This use is similar to the use of (6) in linear network problems.

If some of the network resistors are linear the time required for solution is reduced. This is a result of the fact that neither boundary crossing nor checking for boundary crossing is done for the linear resistors in the network.

We have compared the computation time required to solve identical

TABLE I

Number of Resistors (All Nonlinear)	Time for Crossing a Region (msec)
2	2
5	8
9	20
15	36
24 (11 nodes)	78

problems by this algorithm and the direct iteration method[6] which is given by (11) of Appendix B.

All the networks solved for comparison met the following conditions: (1.) The initial point was a few regions away from the result, (2.) Let ϵ be the number which is compared with the norm of the error after each iteration step to determine termination. In all the tested cases ϵ defined a region which was much smaller than the "typical" regions defined by the characteristics breakpoints. It was found that under these conditions the algorithm performed much better than the iteration method and terminated in a considerably shorter time.

ACKNOWLEDGMENT

The author thanks Drs. J. L. Kelly, Jr., A. J. Goldstein and J. F. Kaiser for reading the manuscript and commenting on it.

APPENDIX A

Crossing of a Boundary Intersection

This appendix considers the case where the solution curve passes through the intersection of two or more boundaries. This case is illustrated in Fig. 5. The problem of finding the region in which the solution curve continues behind the double boundary point arises. Section III suggested finding the region by means of a search procedure. This procedure was judged inefficient since it is lengthy, especially for large networks. The purpose of this appendix is to describe the method by which the algorithm overcomes this difficulty.

Let the boundary point be $\mathbf{e}_{r\lambda_a}$ (Fig. 5). When the solution curve meets a multi-boundary point the algorithm makes a small "jump" across the boundary. The "jump" is a choice of a new point $\mathbf{e}_{rj}$ which is (1) near $\mathbf{e}_{r\lambda_a}$ and (2) nearer (in norm) to the solution than $\mathbf{e}_{r\lambda_a}$ is. The next step is to consider $\mathbf{e}_{rj}$ to be a new initial point and to continue the solution curve from it. The new initial point does not require a new inversion of the conductance matrix but might require one or more successive uses of (6) for obtaining the corresponding matrix inverse.

The point $\mathbf{e}_{rj}$ is found as follows:

$$\mathbf{e}_{rj} = \mathbf{e}_{r\lambda_a} + k(\mathbf{J} - \mathbf{J}_{\lambda a}) \tag{7}$$

where k is given by

$$k = \frac{Min(k_i)}{\| \mathbf{G} \|} \tag{8}$$

where k_i is the smallest slope of the ith resistor and the minimization is carried out on all the resistive tree branches. $\mathbf{G}$ is the conductance matrix of a linear network obtained from the original network by replacing each nonlinear resistor by a linear resistor with a conductance equal to the largest slope of the replaced resistor. $\| \mathbf{G} \|$ denotes the square norm of the matrix.[5] It was proved[6] that (7), written in the form

$$\mathbf{e}_{n+1} = \mathbf{e}_n + k(\mathbf{J} - \mathbf{J}_n) \tag{9}$$

with k given by (8), can be used for solving the nonlinear resistor network problem by iteration. This interation converges in a way which decreases the error in each step and, therefore, the use of (7) implies

$$\| \mathbf{J} - \mathbf{J}_j \| < \| \mathbf{J} - \mathbf{J}_{\lambda_a} \|$$

Thus, the new starting point $\mathbf{e}_{rj}$ is nearer in norm to the solution $\mathbf{e}_r$ than $\mathbf{e}_{r\lambda_a}$; hence the multiple boundary point is passed.

Consider the convergence of a modified algorithm which contains "jumps" whenever the solution curve meets a double boundary point. It will be proved that the modified algorithm converges in a finite number of steps.

Since (7) defines a convergent iteration process it follows, first, that the modified algorithm does converge. The fact that the solution is attained in a finite number of steps is proved in the following way. Let the solution $\mathbf{e}_r$ lie inside some region, say region A. Since the algorithm converges, it will be inside any containing $\mathbf{e}_r$ in a finite number of steps. Once the solution curve is in A, the solution is attained in one step. Thus, the algorithm terminates in a finite number of steps.

In case $\mathbf{e}_r$ is on the boundary, A is considered to consist of all regions which share this boundary. The proof proceeds as above.

APPENDIX B

A Step-By-Step Description of the Algorithm

This section gives a detailed description of the algorithm. It is assumed that the network satisfies conditions (1.) and (2.) of Section II and that a tree τ was chosen such that all voltage sources are tree branches and all current sources are links. The algorithm proceeds as follows:

(1.) Arbitrary values are chosen for the resistor tree branch voltages $\mathbf{e}_{r0}$. Let the region of $\mathcal{E}_r$ in which the point $\mathbf{e}_{r0}$ is located be called region '*a*'.

(2.) Form the conductance matrix $\mathbf{G}_a$ of the linear equivalent network of region 'a' and calculate its inverse $\mathbf{G}_a^{-1}$.

(3.) Calculate the change in the tree voltages

$$\Delta\mathbf{e}_r = \mathbf{G}_a^{-1}(\lambda(\mathbf{J} - \mathbf{J}_0)) \text{ for } \lambda = 1. \tag{10}$$

(4.) Calculate the branch voltages $\mathbf{v}_r$ for all resistors and check if the new voltage value requires the crossing of a boundary.

(5.) Set $\lambda_i = 1$ if no boundary crossing is needed for the ith resistor. If a boundary crossing did occur, set

$$\lambda_i = \frac{v^i - v_{r0}{}^i}{\Delta v_r{}^i}$$

where v^i is the breakpoint voltage of the first boundary that was crossed (Fig. 2), and $\Delta v_{r\lambda}{}^i$ is the change in the ith resistor voltage which corresponds to a change $\Delta\mathbf{e}_r$ in the tree voltages. $v_{r0}{}^i$ is the ith branch voltage which corresponds to $\mathbf{e}_{r0}$.

(6.) Set

$$\lambda_a = \min \lambda_i.$$

This is the largest value of λ for which the point $\mathbf{e}_{r\lambda} = \mathbf{e}_{r0} + \lambda\Delta\mathbf{e}_r$ is in region 'a'. Thus the boundary point of region 'a' is given by

$$\mathbf{e}_{r\lambda_a} = \mathbf{e}_{r0} + \lambda_a\Delta\mathbf{e}_{r\lambda} \tag{11}$$

$$\mathbf{J}_{\lambda_a} = \mathbf{J}_0 + \lambda_a(\mathbf{J} - \mathbf{J}_0). \tag{12}$$

Note that λ_a is the M of Section III.

(7.) If $\lambda_a = 1$, the solution of the problem is in region 'a' and its value is given by (11) and (12). If $\lambda < 1$, the process continues as follows. $\mathbf{e}_{r\lambda_a}$ is on a boundary of region 'a'. This point is either on a boundary between only two regions, as point $\mathbf{e}_{r\lambda_1}$ in Fig. 4, or an intersection of two or more boundaries (Fig. 5). In the former case, the algorithm proceed to Step 8. In the later case the algorithm proceed to Step 9.

(8.) The boundary point $\mathbf{e}_{r\lambda a}$ is on the boundary between regions 'a' and 'b'. The point is considered now as part of the region 'b' and is taken to be the initial point in this region. Thus, $\mathbf{e}_{r0}$ and $\mathbf{J}_0$ are made equal to $\mathbf{e}_{r\lambda a}$ and $\mathbf{J}_{\lambda a}$ respectively. The inverse of the conductance matrix for region 'b' is calculated by modifying the inverse of the conductance matrix for region 'a' and the algorithm return to Step 2 to perform with respect to region 'b' the same operation as was performed with respect to region 'a'.

(9.) Choose a new point $\mathbf{e}_{rj}$ such that

$$\mathbf{e}_{rj} = \mathbf{e}_{r\lambda_a} + k(\mathbf{J} - \mathbf{J}_{\lambda_a})$$

where k is a constant having the dimension of resistance and given by a bound on (8). Consider $\mathbf{e}_{rj}$ to be a new initial point, set the suitable $\mathbf{G}^{-1}$ matrix (by successive use of (6)) and return to Step 3.

REFERENCES

1. Duffin, R. J., Nonlinear Networks IIa, Bull. Am. Math. Soc., *53*, Oct., 1947, pp. 963–971.
2. Birkhoff, G., and Diaz, J. B., Nonlinear Network Problems, Quart. Appl. Math. *13*, Jan., 1956, pp. 431–443.
3. Minty, G. T., Monotone Networks, Proc. Roy. Soc. (London), A, *257*, Sept., 1960, pp. 194–212.
4. Desoer, C. A., and Katzenelson, J., Nonlinear *RLC* Networks, B.S.T.J., *44*, Jan. 1965, pp. 161–198.
5. Faddeeva, V. N., *Computational Methods of Linear Algebra*, Translated by Curtis D. Benster, Chapt. II, Dover Publications, Inc., 1959.
6. Katzenelson, J., and Seitelman, L. H., An Iterative Method for Solving Networks of Nonlinear Resistors, to be published.
7. Minty, G. T., Solving Steady State Nonlinear Networks of 'Monotone' Elements, Trans. IRE. P. G. on Circuit Theory, *CT-8*, 2, June, 1961.
8. Branin, F. H., The Relation between Kron's Method and the Classical Methods of Network Analysis, IRE Wescon Convention Record, Part 2, August, 1959, pp. 3–28.
9. Householder, A. S., A Class of Methods for Inverting Matrices, Trans. Soc. Ind. and Appl. Math. *6*, 1958, pp. 189–195.
10. Householder, A. S., *Principles of Numerical Analysis*, McGraw-Hill, Inc., 1953, p. 79.

Optimum Reception of Binary Sure and Gaussian Signals

By T. T. KADOTA

(Manuscript received May 25, 1965)

The problem of optimum reception of binary sure and Gaussian signals is to specify, in terms of the received waveform, a scheme for deciding between two alternative mean and covariance functions with minimum error probability. In the context of a general treatment of the problem, this article presents a solution which is both mathematically rigorous and convenient for physical application. The optimum decision scheme obtained consists in comparing, with a predetermined threshold c, the sum of a linear and a quadratic form in the received waveform $x(t)$; namely, choose $m_0(t)$ and $r_0(s,t)$ if

$$2\int x(t)g(t)\,dt + \int\int [x(s) - m_1(s)]h(s,t)[x(t) - m_1(t)]\,ds\,dt < c,$$

choose $m_1(t)$ and $r_1(s,t)$ if otherwise, where $m_0(t)$, $m_1(t)$, $r_0(s,t)$ and $r_1(s,t)$ are the two mean and covariance functions, and $g(t)$ is the square-integrable solution of

$$\int r_0(s,t)g(s)\,ds = m_1(t) - m_0(t),$$

while $h(s,t)$ is the symmetric and square-integrable solution of

$$\int\int r_0(s,u)h(u,v)r_1(v,t)\,du\,dv = r_1(s,t) - r_0(s,t).$$

Note that under the assumption of zero mean functions, i.e., $m_0(t) = m_1(t) = 0$, the above result is reduced to the one in a previous article by this author, while with the assumption of identical covariance functions, i.e., $r_0(s,t) = r_1(s,t)$, it is reduced to the classical result essentially obtained by Grenander.

Sections I and II introduce the problem and summarize the main results with certain pertinent remarks, while a detailed mathematical treatment is given in Section III. Although Appendices A–D are not directly required

for solution of the problem, they are added to provide a tutorial background for the results on equivalence and singularity of two Gaussian measures obtained by Grenander, Root and Pitcher as well as some generalization of their results.

I. INTRODUCTION

Suppose the received waveform $x(t)$ observed during the interval $0 \leqq t \leqq 1$ is the sample function of a Gaussian process, whose mean and covariance functions are either $m_0(t)$ and $r_0(s,t)$ or $m_1(t)$ and $r_1(s,t)$. We assume that $m_0(t)$ and $m_1(t)$ are continuous while $r_0(s,t)$ and $r_1(s,t)$ are positive-definite as well as continuous. Denote by H_k, $k = 0,1$, the hypothesis that $m_k(t)$ and $r_k(s,t)$ are the mean and covariance functions of the Gaussian process $\{x_t, 0 \leqq t \leqq 1\}$. Suppose further that α, $0 < \alpha < 1$, and $1 - \alpha$ are the *a priori* probabilities associated with the two hypotheses H_0 and H_1 respectively. Then, reception of binary sure and Gaussian signals may be regarded as a problem of deciding between two hypotheses H_0 and H_1 upon observation of the sample function $x(t)$. Thus, the problem of optimum reception of binary sure and Gaussian signals is to specify a decision scheme in terms of $x(t)$ such that its error probability is minimum.*

In the previous article,[1] a general treatment of the problem was made under the assumption that $m_0(t) = m_1(t) = 0$, and several forms of the optimum decision schemes were given under additional conditions with varying degrees of restriction. The following is most restrictive but most convenient for physical application:

$$\begin{aligned} &\text{choose} \quad H_0 \quad \text{if} \quad \int_0^1 \int_0^1 x(s)h(s,t)x(t)\, ds\, dt < k, \\ &\text{choose} \quad H_1 \quad \text{if} \quad \text{otherwise}, \end{aligned} \tag{1}$$

where $h(s,t)$ is the solution of the integral equation.†

$$\int_0^1 \int_0^1 r_0(s,u)h(u,v)r_1(v,t)\, du\, dv = r_1(s,t) - r_0(s,t) \tag{2}$$

satisfying

$$\int_0^1 \int_0^1 h^2(s,t)\, ds\, dt < \infty, \tag{3}$$

and k is a positive constant (the predetermined threshold); provided the

* A more complete motivation of the problem is given in Ref. 1.

† Existence of such a solution is a part of the condition for (1) to be the optimum decision scheme.

following additional conditions are satisfied for all $i,j = 1,2, \cdots$:

$$a_{ii} > \sum_{j=1}^{\infty}{}' |\, a_{ij} \,|, \quad \frac{\left| \frac{a_{ij}}{\lambda_i} - \delta_{ij} \right|}{a_{jj} - \sum_{k=1}^{\infty}{}' |\, a_{jk} \,|} \leqq K, \tag{4}$$

where λ_i, $i = 1,2, \cdots$, are the eigenvalues of the covariance kernel $r_0(s,t)$ and a_{ij}; $i,j = 1,2, \cdots$, are defined by

$$a_{ij} = \int_0^1 \int_0^1 \psi_i(s) r_1(s,t) \psi_j(t)\, ds\, dt,$$

with $\psi_i(t)$, $i = 1,2, \cdots$, being the orthonormalized eigenfunctions corresponding to λ_i, $i = 1,2, \cdots$, and K is a constant independent of i and j, and the prime above the summation sign signifies omission of the term $j = i$ or $k = j$, whichever the case may be.

As remarked in the previous article, the conditions (4) are not essential to the nature of the problem but are imposed for the sake of mathematical proof. Moreover, they are undesirable from the application viewpoint since they not only are restrictive but also require the explicit knowledge of the eigenvalues and eigenfunctions of the kernel $r_0(s,t)$. Recently, Rao and Varadarajan[2]* and Pitcher[3] have obtained certain general results (on the expression of Radon-Nikodym derivatives), which indicate that such conditions are unnecessary and can be replaced by more meaningful ones. In fact, Rao and Varadarajan extend to the general case where the assumption $m_0(t) = m_1(t) = 0$ is no longer made. The purpose of this article is to generalize the previous results[1] by removing the assumption $m_0(t) = m_1(t) = 0$ and replacing the conditions (4) with more appropriate ones. The first half of the development is a direct generalization of the former Solution — I (the "sampling" approach), while the second half is the application of the results of Grenander[4] and Pitcher to the problem of optimum reception.†

II. SUMMARY AND DISCUSSION OF MAIN RESULTS

As previously stated,[1] the foundation for solution of the problem of optimum reception consists of the following (measure theoretical) facts:

* This article appeared even before the author's previous one,[1] although the current result as well as the previous one were obtained independently.

† The results of Grenander and Pitcher are better suited for this problem than those of Rao and Varadarajan since the former readily yield a concrete specification of the optimum decision scheme comparable to (1)–(3). Although the problem stated at the beginning is solved by a particular combination of Grenander's and Pitcher's result, we have added in appendices an extension of Pitcher's results on equivalence and singularity of two Gaussian measures to the general case where $m_0(t) \neq 0 \neq m_1(t)$ for its own interest.

(1.) Two Gaussian probability measures P_0 and P_1, corresponding respectively to $m_0(t)$ and $r_0(s,t)$ and to $m_1(t)$ and $r_1(s,t)$, can be either "equivalent" or "singular".

(2.) If P_0 and P_1 are equivalent, then there is a certain random variable dP_1/dP_0 called the Radon-Nikodym derivative of P_1 with respect to P_0, which is a function of the sample function $x(t)$, and the following decision scheme yields the minimum non-zero error probability:

$$\begin{aligned} &\text{choose} \quad H_0 \quad \text{if} \quad \frac{dP_1}{dP_0}(x) < \frac{\alpha}{1-\alpha}, \\ &\text{choose} \quad H_1 \quad \text{if} \quad \text{otherwise.} \end{aligned} \tag{5}$$

On the other hand, if P_0 and P_1 are singular, then there is a set N of sample functions such that $P_0(N) = 0$ and $P_1(N) = 1$, thus the error probability of the following decision scheme:

$$\begin{aligned} &\text{choose} \quad H_0 \quad \text{if} \quad x(t) \text{ does not belong to } N, \\ &\text{choose} \quad H_1 \quad \text{if} \quad \text{otherwise,} \end{aligned} \tag{6}$$

is exactly zero, regardless of the *a priori* probabilities, thus resulting in the case of "perfect reception".

Hence, the problem of specifying the optimum decision scheme becomes the problem of finding such a random variable dP_1/dP_0 and a set N as well as a criterion to tell whether P_0 and P_1 are equivalent or singular.

2.1 *Solutions — I*

Suppose $x(t_1), \cdots, x(t_n)$, $0 \leqq t_1 < \cdots < t_n \leqq 1$, are the values of the sample function (the received waveform) sampled at $t_1, \cdots, t_n$, where each sampling interval is to become infinitesimal as $n \to \infty$. Likewise, let $m_0(t_1), \cdots, m_0(t_n)$ be the sampled values of $m_0(t)$. Then, the joint probability density functions for $x(t_1) - m_0(t_1), \cdots, x(t_n) - m_0(t_n)$ under the two hypotheses H_0 and H_1 are obtained by using the mean and ovariance functions $m_0(t)$ and $r_0(s,t)$ (under H_0) and $m_1(t)$ and $r_1(s,t)$ (under H_1).* Then, by forming the ratio of the two density functions, the likelihood ratio l_n of $x(t_1) - m_0(t_1), \cdots, x(t_n) - m_0(t_n)$ is obtained as follows:

* A rather artificial choice of $x(t_i) - m_0(t_i)$, instead of $x(t_i)$, $i = 1, \cdots, n$, is purely for a notational convenience later, and other choices are equally acceptable at this point.

$$l_n(x) = |R_0^{(n)}(R_1^{(n)})^{-1}|^{\frac{1}{2}} \exp\left[\frac{1}{2}\sum_{i,j=1}^{n}[x_{t_i} - m_0(t_i)] \times [(R_0^{(n)})^{-1} - (R_1^{(n)})^{-1}]_{ij}[x_{t_j} - m_0(t_j)] + \sum_{i,j=1}^{n}\left(x_{t_i} - \frac{m_0(t_i) + m_1(t_i)}{2}\right)[(R_1^{(n)})^{-1}]_{ij} \times [m_1(t_j) - m_0(t_j)]\right], \tag{7}$$

where $R_k^{(n)}$, $k = 0, 1$, are $n \times n$ covariance matrices defined by

$$(R_k^{(n)})_{ij} = r_k(t_i, t_j), \qquad k = 0, 1; \qquad i,j = 1, \cdots, n.$$

Next, through the use of martingale theory, the following facts can be established:

P_0 and P_1 are equivalent (the case of non-perfect reception), if and only if

$$\lim_{n\to\infty} |\operatorname{tr}[R_0^{(n)}(R_1^{(n)})^{-1} - 2I + R_1^{(n)}(R_0^{(n)})^{-1} + (R_0^{(n)})^{-1}M^{(n)} + (R_1^{(n)})^{-1}M^{(n)}]| < \infty, \tag{8}$$

where $(M^{(n)})_{ij} = m_i m_j$; $i,j = 1, \cdots, n$,* and m_i, $i = 1, \cdots, n$, are given by

$$m_i = m_1(t_i) - m_0(t_i).$$

In this case

$$\lim_{n\to\infty} l_n(x) = \frac{dP_1}{dP_0}(x) \tag{9}$$

for almost all sample functions under both hypotheses H_0 and H_1.

P_0 and P_1 are singular (the case of perfect reception) if and only if (8) is not satisfied.† In this case, for almost all sample functions,

$$\lim_{n\to\infty} l_n(x) = \begin{cases} 0 \text{ under } H_0, \\ \infty \text{ under } H_1. \end{cases}$$

That is, (8) is a necessary and sufficient condition for the perfect reception to be impossible. The crucial random variable dP_1/dP_0, by which the optimum decision scheme is specified in this case, can be expressed as the limit of the likelihood ratio $l_n(x)$ for almost all sample functions

* "tr" denotes "trace", and I is the $n \times n$ identity matrix.

† In this case, the left-hand side of (8) becomes $+\infty$ necessarily.

$x(t)$. Likewise, negation of (8) is a necessary and sufficient condition for the perfect reception to be possible, and the critical set N can be specified as the set of all sample functions for which the limit of the likelihood ratio is not smaller than any positive constant, say $\alpha/(1-\alpha)$. Therefore, we conclude, in conjunction with (5) and (6), that irrespective of whether or not the condition (8) is satisfied, the optimum decision scheme can be specified as follows:

$$\begin{aligned} &\text{choose} \quad H_0 \quad \text{if} \quad \lim_{n\to\infty} l_n(x) < \frac{\alpha}{1-\alpha}, \\ &\text{choose} \quad H_1 \quad \text{if} \quad \text{otherwise.} \end{aligned} \tag{10}$$

We note in (8) that, if $m_0(t) = m_1(t) = 0$, the trace of the last two terms in the bracket vanishes, thus the necessary and sufficient condition for equivalence of P_0 and P_1 is reduced to

$$\lim_{n\to\infty} \operatorname{tr}[R_0^{(n)}(R_1^{(n)})^{-1} - 2I + R_1^{(n)}(R_0^{(n)})^{-1}] < \infty, \tag{11}$$

which agrees with the previous result.[1] Similarly, if $r_0(s,t) = r_1(s,t) = r(s,t)$, the trace of the first three terms vanishes and the necessary and sufficient condition is reduced to

$$\lim_{n\to\infty} \operatorname{tr}[(R^{(n)})^{-1} M^{(n)}] \equiv \lim_{n\to\infty} (m^{(n)}, (R^{(n)})^{-1} m^{(n)}) < \infty,$$

where $m^{(n)} = (m_i, \cdots, m_n)$.

Now, since the trace of the last two terms in the bracket of (8) is always positive as indicated above, (11) is a necessary condition for (8). Also, since the left-hand side of (11) is known to be either finite or $+\infty$, the conditions

$$\lim_{n\to\infty} \operatorname{tr}[(R_k^{(n)})^{-1} M^{(n)}] < \infty, \qquad k = 0,1 \tag{12}$$

are necessary for (8). Thus, we conclude that a necessary and sufficient condition for equivalence of P_0 and P_1 is that P_0 and P_1 be equivalent in the following three special cases:

(i) $m_0(t) = m_1(t) = 0$,
(ii) $r_0(s,t)$ is substitued for $r_1(s,t)$,
(iii) $r_1(s,t)$ is substituted for $r_0(s,t)$.*

* It can easily be shown that the cases (ii) and (iii) can be combined to the case (iv) where $r_0(s,t) + r_1(s,t)$ is substituted for both $r_0(s,t)$ and $r_1(s,t)$. Thus, the necessary and sufficient condition for equivalence of P_0 and P_1 becomes that they be equivalent in the special cases (i) and (iv). This condition has already been reported elsewhere.[2,3] Furthermore, as it turns out, either the case (ii) or the case (iii) is redundant. That is, P_0 and P_1 are equivalent in general if they are so either in the special cases (i) and (ii) or in (i) and (iii), as shown in Appendix D.

It may be illuminating to rephrase this in terms of the perfect reception of binary (sure and Gaussian) signals, though the use of terms is slightly inconsistent with the remainder of this article. Suppose we consider $m_0(t)$ and $m_1(t)$ as binary sure signals and $r_0(s,t)$ and $r_1(s,t)$ as the covariance functions of binary Gaussian signals or noise whichever the case may be. Then, the perfect reception of the binary sure and Gaussian signals is possible if any one of the following three conditions is satisfied by the constituent signals and noise:

(i′) the perfect reception is possible between the two Gaussian signals alone,

(ii′) the perfect reception of the binary sure signals is possible in the presence of the Gaussian noise with the covariance function $r_0(s,t)$.

(iii′) the condition identical to (ii′) except for $r_0(s,t)$ being replaced by $r_1(s,t)$.

Examination of the form of the likelihood ratio l_n in (7) in conjunction with the decision scheme (10) indicates that, if the exponent and the factor before the exponential converge separately, (10) can be rewritten in terms of their limits. Namely, if there exist a positive constant β and a random variable θ such that

$$\beta = \lim_{n\to\infty} | R_0^{(n)}(R_1^{(n)})^{-1} | , \tag{13}$$

and

$$\begin{aligned}\theta(x) = \lim_{n\to\infty} \Bigg[& \sum_{i,j=1}^{n} [x_{t_i} - m_0(t_i)] [(R_0^{(n)})^{-1} - (R_1^{(n)})^{-1}]_{ij} \\ & \times [x_{t_j} - m_0(t_j)] + 2 \sum_{i,j=1}^{n} \left(x_{t_i} - \frac{m_0(t_i) + m_1(t_i)}{2} \right) \\ & \times [(R_1^{(n)})^{-1}]_{ij} [m_1(t_j) - m_0(t_j)] \Bigg] ,\end{aligned} \tag{14}$$

for almost all sample functions under both hypotheses H_0 and H_1, then (10) is reduced to the following:

$$\begin{aligned} &\text{choose} \quad H_0 \quad \text{if} \quad \theta(x) < \log \left[\frac{1}{\beta} \left(\frac{\alpha}{1-\alpha} \right)^2 \right] , \\ &\text{choose} \quad H_1 \quad \text{if otherwise.} \end{aligned} \tag{15}$$

It can be shown that such β and θ exist if and only if

$$\begin{aligned} &\lim_{n\to\infty} | \operatorname{tr}[R_0^{(n)}(R_1^{(n)})^{-1} - I + (R_1^{(n)})^{-1} M^{(n)}] | < \infty , \\ &\lim_{n\to\infty} | \operatorname{tr}[R_1^{(n)}(R_0^{(n)})^{-1} - I + (R_0^{(n)})^{-1} M^{(n)}] | < \infty . \end{aligned} \tag{16}$$

Note that the above implies the condition (8) as it should. In fact the condition (16) requires not only that the sum of two traces should converge but also that the two traces should converge individually. As we have observed earlier, the condition (8) is equivalent to those of (11) and (12). Hence, the portion of the condition (16) which is additional to (8) is

$$\begin{aligned} &\lim_{n\to\infty} | \operatorname{tr}[R_0^{(n)}(R_1^{(n)})^{-1} - I] | < \infty, \\ &\lim_{n\to\infty} | \operatorname{tr}[R_1^{(n)}(R_0^{(n)})^{-1} - I] | < \infty. \end{aligned} \tag{17}$$

But, according to the previous result,[1] (17) is the necessary and sufficient condition for existence of β and θ when $m_0(t) = m_1(t) = 0$. This is no surprise. For, according to (9), $l_n(x)$ converges for almost all sample functions under both hypotheses when the condition (8) is satisfied. Hence, if in addition the factor before the exponential converges, the exponential must also converge (for almost all sample functions). Thus, the additional condition required is the convergence condition of the factor before the exponential alone. But this factor is obviously independent of $m_0(t)$ and $m_1(t)$. In summary therefore, if and only if the conditions (12) and (17) are satisfied, there exist such β and θ as defined by (13) and (14) and the optimum decision scheme can be specified by (15).

Although the decision scheme (15) is certainly simpler than (10), it is still inconvenient for physical application since it requires the limit operation for each received waveform. What is highly desirable is to express θ of (14) not in terms of the infinite sum but in terms of integrals involving $x(t)$ explicitly. It is completely possible to achieve this objective through a straightforward generalization of Solutions — II of the previous article[1] by removing the assumption $m_0(t) = m_1(t) = 0$.* But, as we have remarked in the Introduction, this method cannot avoid the undesirable accompanying conditions analogous to (4). Hence, in the next subsection, we shall obtain the expression of dP_1/dP_0 directly through a particular combination of the results of Grenander and Pitcher.

2.2 *Solutions — II*

Let us introduce a third Gaussian probability measure P_{10} corresponding to $m_1(t)$ and $r_0(s,t)$. Then, just as equivalence of P_0 and P_1 implies existence of a random variable dP_1/dP_0 (the Radon-Nikodym deriva-

* This generalization has been carried out in detail and the result is contained in an unpublished article by this author.

tive of P_1 with respect to P_0) as stated at the beginning of Section II, equivalence of P_0 and P_{10} and equivalence of P_{10} and P_1 imply existence of dP_{10}/dP_0 and dP_1/dP_{10} respectively. Now recall that the key to the solution is to find an expression of dP_1/dP_0 in terms of $x(t)$, in the case where the condition (8) is satisfied. Note that the term, Radon-Nikodym *derivative*, and its symbol immediately suggest the following formalism which is analogous to the chain rule in calculus:

$$\frac{dP_1}{dP_0} = \frac{dP_1}{dP_{10}} \frac{dP_{10}}{dP_0}. \tag{18}$$

According to measure theory, P_0 and P_1 are equivalent and (18) is valid for almost all sample functions under the hypotheses H_0, H_{10} and H_1,* if P_0 and P_{10} as well as P_{10} and P_1 are equivalent. Thus, the task of finding an expression for dP_1/dP_0 in terms of $x(t)$ is equivalent to that of finding such expressions for dP_{10}/dP_0 and dP_1/dP_{10} together with the conditions for equivalence.

Now, through the application of the condition (8) to the case of two Gaussian measures P_0 and P_{10}, it is seen that P_0 and P_{10} are equivalent if and only if (12) with $k = 0$ is satisfied. Note that this is the special case (ii) in the preceding subsection, namely, that perfect reception of the binary sure signals $m_0(t)$ and $m_1(t)$ is not possible in the presence of Gaussian noise with the covariance function $r_0(s,t)$. Then, according to Grenander,[4] if the integral equation

$$\int_0^1 r_0(s,t)g(s)ds = m_1(t) - m_0(t), \, 0 \leqq t \leqq 1, \tag{19}$$

has a square-integrable solution $g(t)$, then dP_{10}/dP_0 can be expressed as

$$\frac{dP_{10}}{dP_0}(x) = \exp\left\{\int_0^1 \left[x(t) - \frac{m_0(t) + m_1(t)}{2}\right] g(t)\, dt\right\} \tag{20}$$

for almost all sample functions under the hypotheses H_0 and H_{10}. As we may recall, it is through the substitution of (20) into (5) that the well-known optimum receiver (decision scheme) of binary sure signals in noise is obtained; namely,

choose H_0 if

$$\int_0^1 x(t)g(t)\, dt < \frac{1}{2}\int_0^1 [m_0(t) + m_1(t)]g(t)\, dt + \log\frac{\alpha}{1-\alpha}, \tag{21}$$

choose H_{10} if otherwise.

* H_{10} is the hypothesis that $m_1(t)$ and $r_0(s,t)$ are the mean and covariance functions of the Gaussian process $\{x_t, 0 \leqq t \leqq 1\}$.

Similarly, from the condition (8), two Gaussian measures P_{10} and P_1 are equivalent if and only if (11) is satisfied. This is essentially equal to the special case (i), namely, that the perfect reception is not possible between two Gaussian signals with $r_0(s,t)$ and $r_1(s,t)$, where $x(t) - m_1(t)$ instead of $x(t)$ is to be regarded as the sample function in this case. Then, according to the previous result,[1] which is improved by Pitcher,[3] if the integral equation (2) has a solution $h(s,t)$ which is symmetric and satisfies (3), dP_1/dP_{10} can be expressed as

$$\frac{dP_1}{dP_{10}}(x) = \left(\prod_{i=1}^{\infty} \rho_i\right)^{-\frac{1}{2}} \exp\left\{\frac{1}{2}\int_0^1 \int_0^1 [x(s) - m_1(s)]h(s,t)[x(t) - m_1(t)]\, ds\, dt\right\} \tag{22}$$

for almost all sample functions under the hypotheses H_{10} and H_1, where $\rho_i > 0$, $i = 1, 2, \cdots$, are the eigenvalues of a certain operator defined in terms of $r_0(s,t)$ and $r_1(s,t)$. As in the preceding case, it is seen that substitution of (22) into (5) yields the optimum decision scheme (1).

In summary, therefore, if the integral equations (19) and (2) have a square-integrable solution $g(t)$ and a symmetric and square-integrable (in the sense of (3)) solution $h(s,t)$ respectively, then the crucial random variable dP_1/dP_0 can be expressed as the product of the right-hand sides of (20) and (22) for almost all sample functions under H_0 and H_1. Thus, the desired optimum decision scheme becomes the following:

choose H_0 if

$$2\int_0^1 x(t)g(t) + \int_0^1 \int_0^1 [x(s) - m_1(s)]h(s,t)[x(t) - m_1(t)]\, ds\, dt < \int_0^1 [m_0(t) + m_1(t)]g(t)dt + \log\left[\frac{1}{\hat{\mathfrak{B}}}\left(\frac{\alpha}{1-\alpha}\right)^2\right], \tag{23}$$

choose H_1 if otherwise,

where

$$\hat{\mathfrak{B}}^{-1} = \prod_{i=1}^{\infty} \rho_i .$$

It should be remarked that the indices 0 and 1 can be consistently interchanged throughout. This follows from the symmetry of the problem with respect to the indices. Moreover, by virtue of the symmetry of $h(s,t)$ in s and t, the indices on the left-hand side of (2) can be inter-

changed while the right-hand side remains unchanged. We also remark that the solutions $g(t)$ and $h(s,t)$ of the integral equations (19) and (2) respectively are unique under the constraints of square-integrability for $g(t)$ and symmetry and square-integrability in the sense of (3) for $h(s,t)$.

Physical interpretation of the optimum decision scheme (23) is obvious, at least, in principle. Given two alternative mean and covariance functions $m_0(t)$ and $r_0(s,t)$, and $m_1(t)$ and $r_1(s,t)$, the optimum receiver consists of a linear and a quadratic filter whose impulse responses are $g(t)$ and $h(s,t)$, respectively, and whose inputs are $2x(t)$ and $x(t) - m_1(t)$ respectively. The outputs of the two filters are sampled at the end of the observation interval, and the decision is made by comparing the sum of the two sampled outputs with the predetermined threshold c, namely, the right-hand side of the inequality in (23).

Finally, although somewhat redundant, it seems instructive to examine the optimum decision scheme in the two special cases which have already been considered.

Case 1:

$$r_0(s,t) = r_1(s,t) = r(s,t),$$

namely, the case of reception of binary sure signals $m_0(t)$ and $m_1(t)$ in the presence of Gaussian noise with the covariance function $r(s,t)$. In this case, the second integral in the inequality of the optimum decision scheme (23) vanishes, since the right-hand side of the integral equation (2) becomes identically zero, thus yielding the identically vanishing function as the only solution satisfying the conditions of symmetry and square-integrability (3), i.e., $h(s,t) = 0$. Moreover, $\hat{\beta}$ becomes unity since all ρ_i, $i = 1,2, \cdots$, are unity. Hence, the optimum decision scheme (23) is reduced to that of (21) where $g(t)$ is the square-integrable solution of (19) with $r_0(s,t)$ replaced by $r(s,t)$.

Case 2:

$$m_0(t) = m_1(t) = 0,$$

namely, the case of reception of binary Gaussian signals with the covariance functions $r_0(s,t)$ and $r_1(s,t)$. In this case, the first and the third integrals in the inequality of (23) vanish, since the right-hand side of the integral equation (19) becomes identically zero, thus admitting the trivial solution as the only square-integrable solution, i.e., $g(t) = 0$. Hence, the optimum decision scheme (23) is reduced essentially to (1).

III. MATHEMATICAL THEORY

3.1 *Statement of Problem*

Definitions

Let Ω be the space of all real-valued functions $\omega(\cdot)$ on [0,1], and let $\tilde{x}_t(\cdot)$ be a real-valued function defined on Ω such that the value of $\tilde{x}_t(\cdot)$ at ω is equal to $\omega(t)$. Let $\tilde{\mathfrak{B}}$ be the σ-field generated by the class of all sets of the form

$$\{\omega: (\tilde{x}_{t_1}(\omega), \cdots, \tilde{x}_{t_n}(\omega)) \in A\}, \tag{24}$$

where n and $t_i \in [0,1]$, $i = 1, \cdots, n$ are arbitrary and A is any n-dimensional Borel set. Finally, let $\tilde{P}_0$, $\tilde{P}_{10}$ and $\tilde{P}_1$ be Gaussian measures induced on $\tilde{\mathfrak{B}}$ respectively by m_0 and r_0, by m_1 and r_0, and by m_1 and r_1, where m_k, $k = 0,1$, are real-valued, continuous functions on [0,1], while r_k, $k = 0,1$, are real-valued, symmetric, positive-definite, continuous functions on $[0,1] \times [0,1]$.* Then, $\tilde{x}_t$ is obviously $\tilde{\mathfrak{B}}$-measurable for every $t \in [0,1]$, thus $\{\tilde{x}_t, 0 \leqq t \leqq 1\}$ is a real Gaussian process whose finite dimensional distributions are given by the values of $\tilde{P}_0$, $\tilde{P}_{10}$ and $\tilde{P}_1$ on the set defined by (24). Since m_k and r_k, $k = 0,1$, are continuous, there always exists a separable (with respect to all $\tilde{P}_0$, $\tilde{P}_{10}$ and $\tilde{P}_1$) and measurable version of $\{\tilde{x}_t, 0 \leqq t \leqq 1\}$, which we denote by $\{x_t, 0 \leqq t \leqq 1\}$.† Let $\mathfrak{B}$ be the minimal σ-field with respect to which x_t is measurable for every $t \in [0,1]$, and let P_0, P_{10} and P_1 be the restrictions of $\tilde{P}_0$, $\tilde{P}_{10}$ and $\tilde{P}_1$ respectively on $\mathfrak{B}$.

Next, define a set function $P_\alpha(\Lambda)$, $0 < \alpha < 1$, $\Lambda \in \mathfrak{B}$, by

$$P_\alpha(\Lambda) = \alpha P_0(\Lambda) + (1 - \alpha)P_1(\Omega - \Lambda).$$

Let Λ_α be such a set that

$$P_\alpha(\Lambda_\alpha) \leqq P_\alpha(\Lambda) \quad \text{for all} \quad \Lambda \in \mathfrak{B}.$$

Problem

Given α, $0 < \alpha < 1$, specify such a set Λ_α in terms of x_t.

* See Ref. 5, pp. 609–610 and p. 72.

† Let $\tilde{P}$ be a probability measure on $\tilde{\mathfrak{B}}$ with respect to which all $\tilde{P}_0$, $\tilde{P}_{10}$ and $\tilde{P}_1$ are absolutely continuous, e.g., $\tilde{P} = \frac{1}{3}(\tilde{P}_0 + \tilde{P}_{10} + \tilde{P}_1)$. Now continuity of m_k and r_k, $k = 0,1$, implies continuity in probability of $\tilde{x}_t$ on [0,1] with respect to $\tilde{P}_0$, $\tilde{P}_{10}$ and $\tilde{P}_1$, hence with respect to $\tilde{P}$. Then, there exists a separable (with respect to $\tilde{P}$) and measurable version of $\{\tilde{x}_t, 0 \leqq t \leqq 1\}$, (see Ref. 5, pp. 54–59). But, because $\tilde{P}_0$, $\tilde{P}_{10}$, $\tilde{P}_1 \ll \tilde{P}$, the same version is separable with respect to $\tilde{P}_0$, $\tilde{P}_{10}$ and $\tilde{P}_1$ also.

3.2 *Solution*

Preliminaries

The foundation for solving the above problem consists of the following two measure theoretical facts:

(a.) The Gaussian measures P_0 and P_1 can be either equivalent, $P_0 \equiv P_1$, or singular, $P_0 \perp P_1$.[2,6,7,8,9] *

$$\text{(b.)} \qquad \text{If } P_0 \equiv P_1, \quad \text{then} \quad \Lambda_\alpha = \left\{\omega: \frac{dP_1}{dP_0}(\omega) \geqq \frac{\alpha}{1-\alpha}\right\}, \tag{25}$$

$$\text{if } P_0 \perp P_1, \quad \text{then} \quad \Lambda_\alpha = N,$$

where dP_1/dP_0 is the Radon-Nikodym derivative of P_1 with respect to P_0 and N is a $\mathcal{B}$-measurable set such that $P_0(N) = 0 = P_1(\Omega - N)$.[1]

Thus, the problem stated in the preceding subsection is reduced to that of finding dP_1/dP_0 if $P_0 \equiv P_1$ and N is $P_0 \perp P_1$, which are expressible in terms of x_t.

Solutions — I

Let $\{\tau_k\}$ be a sequence of points in [0,1], which is dense in [0,1]. Let $\mathcal{B}_n$ be the minimal σ-field with respect to which all x_{τ_i}, $i = 1, \cdots, n$, are measurable, and let $\mathcal{B}_\infty$ be the minimal σ-field containing $\bigcup_{n=1}^{\infty} \mathcal{B}_n$. Obviously,

$$\mathcal{B}_1 \subset \mathcal{B}_2 \subset \cdots \subset \mathcal{B}_\infty \subset \mathcal{B}. \tag{26}$$

Then, since $\{x_t, 0 \leqq t \leqq 1\}$ is continuous in probability (with respect to P_0), it follows that, for every set $\Lambda \in \mathcal{B}$, there exists a set $\Lambda' \in \mathcal{B}_\infty$ such that

$$P_0(\Lambda \Delta \Lambda') = 0. \tag{27}$$

Now, from the fact that m_k and r_k, $k = 0,1$, are two alternative mean and covariance functions of $\{x_t, 0 \leqq t \leqq 1\}$, the density functions p_0 and p_1 of the random variables $x_{\tau_i}(\omega) - m_0(\tau_i)$, $i = 1, \cdots, n$, corresponding to P_0 and P_1 respectively, are obtained as follows:

* Also see Theorem 3 in Appendix D.

$$p_0(\nu_1, \cdots, \nu_n) = (2\pi)^{-n/2} \mid R_0^{(n)} \mid^{-\frac{1}{2}} \exp\left[-\frac{1}{2}\sum_{i,j=1}^{n} \nu_i[(R_0^{(n)})^{-1}]_{ij}\,\nu_j\right],$$

$$p_1(\nu_1, \cdots, \nu_n) = (2\pi)^{-n/2} \mid R_1^{(n)} \mid^{-\frac{1}{2}} \exp \left[-\frac{1}{2}\sum_{i,j=1}^{n} (\nu_i - m_i)\,[(R_1^{(n)})^{-1}]_{ij}(\nu_j - m_j)\right],$$

where $R_k^{(n)}$, $k = 0,1$, are $n \times n$ symmetric, positive-definite matrices defined by

$$(R_k^{(n)})_{ij} = r_k(\tau_i, \tau_j), \qquad k = 0, 1; \qquad i,j = 1, \cdots, n,$$

and

$$m_i = m_1(\tau_i) - m_0(\tau_i), \qquad i = 1, \cdots, n.$$

Then, define a random variable l_n by

$$\begin{aligned} l_n(\omega) &= \frac{p_1[x_{\tau_1}(\omega) - m_0(\tau_1), \cdots, x_{\tau_n}(\omega) - m_0(\tau_n)]}{p_0[x_{\tau_1}(\omega) - m_0(\tau_1), \cdots, x_{\tau_n}(\omega) - m_0(\tau_n)]} \\ &= \mid R_0^{(n)}(R_1^{(n)})^{-1} \mid^{\frac{1}{2}} \exp\left[\frac{1}{2}\sum_{i,j=1}^{n} [x_{\tau_i}(\omega) - m_0(\tau_i)]\right. \\ &\quad \times [(R_0^{(n)})^{-1} - (R_1^{(n)})^{-1}]_{ij}\,[x_{\tau_j}(\omega) - m_0(\tau_j)] \\ &\quad + \sum_{i,j=1}^{n}\left[x_{\tau_i}(\omega) - \frac{m_0(\tau_i) + m_1(\tau_i)}{2}\right][(R_1^{(n)})^{-1}]_{ij} \\ &\quad \left.\times\,[m_1(\tau_j) - m_0(\tau_j)]\right]. \end{aligned} \tag{28}$$

Note that $l_n(\omega) \geqq 0$ for all n, and $p_1 = 0$ whenever $p_0 = 0$ and vice versa. Hence, the processes $\{l_n, n \geqq 1\}$ and $\{1/l_n, n \geqq 1\}$ are martingales with respect to P_0 and P_1 respectively.* Then, $\lim_{n\to\infty} l_n$ exists a.e. (P_0) and is denoted by l_∞, and also $\lim_{n\to\infty} 1/l_n$ exists a.e. (P_1).† Furthermore, it can be shown that

(i) if $P_0 \equiv P_1$, then (26) and (27) imply

$$l_\infty = \frac{dP_1}{dP_0}, \quad \text{a.e. } (P_0), \tag{29}$$

(ii) if $P_0 \perp P_1$, then

$$P_0(\{\omega: \lim_{n\to\infty} l_n(\omega) \geqq c\}) = 0 = P_1(\{\omega: \lim l_n(\omega) < c\}) \tag{30}$$

for an arbitrary constant $c > 0$.‡

* See Ref. 5, pp. 91–93.
† See Ref. 5, p. 319.
‡ See Ref. 1, pp. 2783–2784. Although the definition of l_n is slightly different from the one in Ref. 1 the derivation procedure is identical.

Thus, upon combination of (29) and (30) in conjunction with (25), the desired set Λ_α can be given by

$$\Lambda_\alpha = \left\{\omega: \lim_{n\to\infty} l_n(\omega) \geqq \frac{\alpha}{1-\alpha}\right\},$$

irrespective of whether $P_0 \equiv P_1$ or $P_0 \perp P_1$.

Under certain restrictive conditions, the set Λ_α can be specified in terms of well defined functions of x_t. It is of interest to obtain such specifications as well as the accompanying conditions in terms of the given mean and covariance functions m_k and r_k, $k = 0,1$.

If $P_0 \equiv P_1$, it has already been shown that

$$\Lambda_\alpha = \left\{\omega: l_\infty(\omega) \geqq \frac{\alpha}{1-\alpha}\right\}.$$

Furthermore, it can be shown through the use of martingale theory that $P_0 \equiv P_1$ if and only if (8) is satisfied.*

Next, examination of (28) indicates that, in addition to the condition (8), if there exists a positive constant β such that (13) holds, then there exists a random variable θ such that

$$\begin{aligned}\theta(\omega) = \lim_{n\to\infty}\Bigg[&\sum_{i,j=1}^{n} [x_{\tau_i}(\omega) - m_0(\tau_i)]\,[(R_0^{(n)})^{-1} - (R_1^{(n)})^{-1}]_{ij} \\ &\times [x_{\tau_j}(\omega) - m_0(\tau_j)] + 2\sum_{i,j=1}^{n}\left(x_{\tau_i}(\omega) - \frac{m_0(\tau_i) + m_1(\tau_i)}{2}\right) \\ &\times [(R_1^{(n)})^{-1}]_{ij}\,[m_1(\tau_j) - m_0(\tau_j)]\Bigg], \quad \text{a.e.}(P_0).\end{aligned}$$

Thus, the set Λ_α can be specified as follows:

$$\Lambda_\alpha = \left\{\omega: \theta(\omega) \geqq \log\left[\frac{1}{\beta}\left(\frac{\alpha}{1-\alpha}\right)^2\right]\right\}.$$

It can be shown through the use of martingale theory that the conditions (8) and (13) are equivalent to those of (16).†

Solutions — II

Let R_0 and R_1 be the integral operators whose kernels are r_0 and r_1 respectively, that is, for any real-valued function f,

* See Ref. 1, pp. 2784–2785, with the definition of l_n replaced by (28) of this article.

† See Ref. 1, pp. 2786–2787, with the definition of l_n replaced by (28) of this article.

$$(R_k f)(t) = \int_0^1 r_k(s,t) f(s)\, ds, \qquad 0 \leqq t \leqq 1, \qquad k = 0,1,$$

whenever the right-hand side is well defined. Then, Grenander shows that*

if there exists $g \in \mathfrak{L}_2(0,1)$† satisfying the integral equation (19), then $P_0 \equiv P_{10}$ and

$$\frac{dP_{10}}{dP_0} = \exp\left\{\int_0^1 \left[x_t - \frac{m_0(t) + m_1(t)}{2}\right] g(t)\, dt\right\}, \qquad \text{a.e.}(P_0).$$

On the other hand, according to the previous result, improved by Pitcher,‡

if there exists a symmetric function h on $[0,1] \times [0,1]$ satisfying (3) and the integral equation (2), then $P_{10} \equiv P_1$ and

$$\frac{dP_1}{dP_{10}} = |\, R_0^{-\frac{1}{2}} R_1 R_0^{-\frac{1}{2}} \,|^{-\frac{1}{2}} \exp\left\{\frac{1}{2}\int_0^1 \int_0^1 [x_s - m_1(s)]h(s,t) \times [x_t - m_1(t)]\, ds\, dt\right\}, \qquad \text{a.e.}(P_{10})\S.$$

Since $P_0 \equiv P_{10}$ and $P_{10} \equiv P_1$ imply $P_0 \equiv P_1$ and

$$\frac{dP_1}{dP_0} = \frac{dP_1}{dP_{10}} \frac{dP_{10}}{dP_0}, \qquad \text{a.e.}(P_0),$$

we conclude that

if there exist $g \in \mathfrak{L}_2(0,1)$ satisfying (19) and symmetric h satisfying (3) and (2), then $P_0 \equiv P_1$ and

$$\frac{dP_1}{dP_0} = |\, R_0^{-\frac{1}{2}} R_1 R_0 \,|^{-\frac{1}{2}} \exp\left\{\int_0^1 \left[x_t - \frac{m_0(t) + m_1(t)}{2}\right] g(t)\, dt + \frac{1}{2}\int_0^1 \int_0^1 [x_s - m_1(s)]h(s,t)[x_t - m_1(t)]\, ds\, dt\right\}, \qquad \text{a.e. } (P_0).$$

Therefore, through the substitution of the above into (25), the desired set Λ_α can be specified as follows:

* See Appendix B.

† $\mathfrak{L}_2(0,1)$ is the space of all square-integrable functions on $[0,1]$.

‡ See Appendix C.

§ $|\, R_0^{-\frac{1}{2}} R_1 R_0^{-\frac{1}{2}} \,| = \prod_{i=1}^{\infty} \rho_i$ where ρ_i, $i = 1, 2, \cdots$, are the eigenvalues of $R_0^{-\frac{1}{2}} R_1 R_0^{-\frac{1}{2}}$.

$$\Lambda_\alpha = \Big\{\omega\colon 2\int_0^1 x_t(\omega)g(t)\,dt + \int_0^1\int_0^1 [x_s(\omega) - m_1(s)]h(s,t)[x_t(\omega) - m_1(t)]\,ds\,dt \geqq \int_0^1 [m_0(t) + m_1(t)]g(t)\,dt + \log\left(\frac{\alpha}{1-\alpha}\right)^2 |\, R_0^{-\frac{1}{2}} R_1 R_0^{-\frac{1}{2}} |\Big\},$$

if there exist $g \in \mathfrak{L}_2(0,1)$ satisfying (19) and symmetric h satisfying (3) and (2).

IV. ACKNOWLEDGMENT

The author has greatly benefitted from discussions with T. S. Pitcher (MIT Lincoln Laboratory) and W. L. Root (University of Michigan) as well as with S. P. Lloyd and J. McKenna (Bell Telephone Laboratories, Inc.).

APPENDICES

These appendices are given primarily for a tutorial reason. The majority of the theorems and lemmas here are taken from two articles by Root[9] and Pitcher,[3] in the original, modified or extended forms.* Lemmas 1, 2, and 3 are in the original modified and extended form respectively. Theorem 1 is supplemented by (iii) and a corollary. A more significant supplement, however, is in its proof. While the extended portion of Lemma 4 is a routine matter (hence its proof is omitted), Lemma 5 is significantly extended and strengthened. Lemmas 6 and 7† are added as a supplementary part of the proof of Theorem 2. Although Theorem 2 is stated somewhat differently and in much more detail, its main content remains the same. While the first corollary to Theorem 2 is almost obvious, the proof of the second is considerably involved and is given as "Theorem 3" in Ref. 3. Lemmas 8 and 9 and Theorem 3, which is a generalization of Theorems 1 and 2, are the author's addition. However, their major contents have already been reported elsewhere in different forms, e.g., Ref. 2, including the two corollaries to Theorem 3.

* The term "extended" refers to the extension of the results in Refs. 9 and 3 to the case where the assumption $m_0 = m_1 = 0$ is no longer made.

† The proof of Lemma 7 is supplied by both Root and Pitcher.

APPENDIX A

Preliminaries

Let ρ and μ be probability measures defined on a σ-field $\mathfrak{F}$ of subsets of an infinite set (uncountable in general). Let $\bar{\rho}$ and $\bar{\mu}$ be the completions of ρ and μ on σ-fields $\bar{\mathfrak{F}}_\rho$ and $\bar{\mathfrak{F}}_\mu$ respectively.

Lemma 1: Let $\mathfrak{F}_0$ be a σ-field such that

$$\mathfrak{F}_0 \subset \bar{\mathfrak{F}}_\rho \quad \text{and} \quad \mathfrak{F}_0 \subset \bar{\mathfrak{F}}_\mu ,$$

and let ρ_0 and μ_0 be the restrictions of $\bar{\rho}$ and $\bar{\mu}$ on $\mathfrak{F}_0$. Then,

$$\rho_0 \perp \mu_0 \Rightarrow \rho \perp \mu.$$

Lemma 2: Assume

$$\rho_0 \equiv \mu_0 .$$

Let $\bar{\mathfrak{F}}_0$ be a σ-field of sets of the form $\Lambda \triangle N$, $\Lambda \in \mathfrak{F}_0$, $\bar{\rho}(N) = 0$. Assume

$$\mathfrak{F} \subset \bar{\mathfrak{F}}_0 .$$

Let ρ'_0 and μ'_0 be the restrictions of $\bar{\rho}$ and $\bar{\mu}$ on $\mathfrak{F}_0$, and let ρ' and μ' be the restrictions of ρ'_0 and μ'_0 on $\mathfrak{F}$. Then,

(*i*) $\rho = \rho'$ *and* $\mu = \mu'$,
(*ii*) $\rho \equiv \mu$,
(*iii*) $\bar{\mathfrak{F}}_0 = \bar{\mathfrak{F}}_\rho = \bar{\mathfrak{F}}_\mu$,
(*iv*) $\dfrac{d\mu}{d\rho} = \dfrac{d\mu_0}{d\rho_0}$, *a.e.* (ρ).

Lemma 3: Let $\theta_1, \theta_2, \cdots$, be a sequence of Gaussian variables ($\mathfrak{F}$-measurable) with respect to both ρ and μ such that

$$E_\rho\{\theta_i\} = 0, \qquad E_\mu\{\theta_i\} = \nu_i ,$$

$$E_\rho\{\theta_i\theta_j\} = \alpha_i\delta_{ij} , \qquad E_\mu\{(\theta_i - \nu_i)(\theta_j - \nu_j)\} = \beta_i\delta_{ij} ,$$

where E_ρ and E_μ denote the expectations with respect to ρ and μ respectively and α_i, β_i, $i = 1, 2, \cdots$, are arbitrary positive numbers. Let $\hat{\mathfrak{F}}$ be the minimal σ-field with respect to which all θ_i, $i = 1, 2, \cdots$, are measurable, and let $\hat{\rho}$ and $\hat{\mu}$ be the restrictions of ρ and μ on $\hat{\mathfrak{F}}$. Then,

(*i*) *either* $\hat{\rho} \equiv \hat{\mu}$ *or* $\hat{\rho} \perp \hat{\mu}$,
(*ii*) $\hat{\rho} \equiv \hat{\mu}$ *if and only if*

$$\sum_{i=1}^{\infty} \left(1 - \frac{\alpha_i}{\beta_i}\right)^2 < \infty \quad \text{and} \quad \sum_{i=1}^{\infty} \frac{\nu_i^2}{\alpha_i + \beta_i} < \infty .$$

(*iii*) *if* $\hat{\rho} \equiv \hat{\mu}$,

$$\frac{d\hat{\mu}}{d\hat{\rho}} = \exp\left\{\sum_{i=1}^{\infty}\left[\frac{1}{2}\left(\frac{1}{\alpha_i} - \frac{1}{\beta_i}\right)\theta_i^2 + \frac{\nu_i}{\beta_i}\left(\theta_i - \frac{\nu_i}{2}\right) + \tfrac{1}{2}\log\frac{\alpha_i}{\beta_i}\right]\right\},$$

a.e. $(\hat{\rho})$.

Proof:

(i) Let $\hat{\mathfrak{F}}_i$, $i = 1, 2, \cdots$, be the minimal σ-field with respect to which θ_i is measurable, and let $\hat{\rho}^{(i)}$ and $\hat{\mu}^{(i)}$ be the restrictions of ρ and μ on $\hat{\mathfrak{F}}_i$. Then, from the hypothesis of the lemma,

$$\hat{\rho}^{(i)} \equiv \hat{\mu}^{(i)}, \qquad i = 1, 2, \cdots,$$

and

$$\hat{\mathfrak{F}} = \prod_{i=1}^{\infty} \hat{\mathfrak{F}}_i, \qquad \hat{\rho} = \prod_{i=1}^{\infty} \hat{\rho}^{(i)}, \qquad \hat{\mu} = \prod_{i=1}^{\infty} \hat{\mu}^{(i)}.$$

Hence, the assertion (i) follows from Kakutani's theorem.*

(ii) From the hypothesis of the lemma,

$$\frac{d\hat{\mu}^{(i)}}{d\hat{\rho}^{(i)}} = \left(\frac{\alpha_i}{\beta_i}\right)^{\frac{1}{4}} \exp\left[\frac{1}{2}\left(\frac{1}{\alpha_i} - \frac{1}{\beta_i}\right)\theta_i^2 + \frac{\nu_i}{\beta_i}\theta_i - \frac{\nu_i^2}{2\beta_i}\right], \qquad \text{a.e. } (\rho). \quad (31)$$

Thus,†

$$E_{\hat{\rho}^{(i)}}\left\{\left(\frac{d\hat{\mu}^{(i)}}{d\hat{\rho}^{(i)}}\right)^{\frac{1}{2}}\right\} = \int_{-\infty}^{\infty}\left(\frac{\alpha_i}{\beta_i}\right)^{\frac{1}{4}} \exp\left[\frac{1}{4}\left(\frac{1}{\alpha_i} - \frac{1}{\beta_i}\right)\zeta^2 + \frac{\nu_i}{2\beta_i}\zeta - \frac{\nu_i^2}{4\beta_i}\right]$$

$$\cdot (2\pi\alpha_i)^{-\frac{1}{2}} \exp\left(-\frac{\zeta^2}{2\alpha_i}\right) d\zeta$$

$$= \frac{(4\alpha_i\beta_i)^{\frac{1}{4}}}{(\alpha_i + \beta_i)^{\frac{1}{2}}} \exp\left(-\frac{1}{4}\frac{\nu_i^2}{\alpha_i + \beta_i}\right).$$

Note, for all $i = 1, 2, \cdots$,

$$\frac{4\alpha_i\beta_i}{(\alpha_i + \beta_i)^2} \leqq 1 \quad \text{and} \quad 0 < \exp\left(-\frac{1}{4}\frac{\nu_i^2}{\alpha_i + \beta_i}\right) < 1.$$

Hence

$$\prod_{i=1}^{\infty} E_{\hat{\rho}^{(i)}}\left\{\left(\frac{d\hat{\mu}^{(i)}}{d\hat{\rho}^{(i)}}\right)^{\frac{1}{2}}\right\}$$

converges to a positive number if both

* See Ref. 9, pp. 295–296.

† $E_{\hat{\rho}^{(i)}}$ denotes expectation with respect to $\hat{\rho}^{(i)}$.

$$\prod_{i=1}^{\infty} \frac{4\alpha_i\beta_i}{(\alpha_i + \beta_i)^2} \quad \text{and} \quad \sum_{i=1}^{\infty} \frac{\nu_i^2}{\alpha_i + \beta_i}$$

converge. Therefore, according to Kakutani's theorem, $\hat{\rho} \equiv \hat{\mu}$ if and only if these infinite product and sum converge.
Now,

$$\prod_{i=1}^{\infty} \frac{4\alpha_i\beta_i}{(\alpha_i + \beta_i)^2}$$

converges if and only if*

$$\sum_{i=1}^{\infty} \left[1 - \frac{4\alpha_i\beta_i}{(\alpha_i + \beta_i)^2} \right] < \infty.$$

But

$$1 - \frac{4\alpha_i\beta_i}{(\alpha_i + \beta_i)^2} = \left(1 - \frac{\alpha_i}{\beta_i}\right)^2 \Big/ \left(1 + \frac{\alpha_i}{\beta_i}\right)^2$$

and the infinite sum of this converges if and only if

$$\sum_{i=1}^{\infty} \left(1 - \frac{\alpha_i}{\beta_i}\right)^2 < \infty,$$

hence, the assertion (ii) follows.

(iii) Note

$$\prod_{i=1}^{n} \frac{d\hat{\mu}^{(i)}}{d\hat{\rho}^{(i)}} = E_{\hat{\rho}} \left\{ \frac{d\hat{\mu}}{d\hat{\rho}} \,\middle|\, \prod_{i=1}^{n} \bar{\mathfrak{F}}_i \right\}, \qquad \text{a.e. } (\hat{\rho}).$$

Hence,*

$$\prod_{i=1}^{\infty} \frac{d\hat{\mu}^{(i)}}{d\hat{\rho}^{(i)}} = \frac{d\hat{\mu}}{d\hat{\rho}}, \qquad \text{a.e. } (\hat{\rho}).$$

Then, the assertion (iii) is obtained through substitution of (31) into the above.

APPENDIX B

First Theorem on Equivalence and Singularity

Theorem 1: (*Grenander*)

(*i*) *Either* $P_0 \equiv P_{10}$ *or* $P_0 \perp P_{10}$,
(*ii*) $P_0 \equiv P_{10}$ *if and only if* $R_0^{-\frac{1}{2}} m \in \mathfrak{L}_2(0,1)$,
(*iii*) *if* $P_0 \equiv P_{10}$,

* See Ref. 11, p. 381.
† See Ref. 5, p. 331.

$$\frac{dP_{10}}{dP_0} = \exp\left\{\sum_{i=1}^{\infty}\left[\frac{\nu_i}{\lambda_i}\xi_i - \frac{\nu_i^2}{2\lambda_i}\right]\right\}, \qquad a.e.\ (\bar{P}_0),$$

where $m(t) = m_1(t) - m_0(t)$, $0 \leqq t \leqq 1$, *and* ξ_i *and* ν_i, $i = 1, 2, \cdots$ *are defined by*

$$\xi_i(\omega) = (x(\omega) - m_0, \psi_1) \equiv \int_0^1 [x_t(\omega) - m_0(t)]\psi_i(t)\, dt,$$

$$a.e.\ (\bar{P}_0, \bar{P}_{10})$$

$$\nu_i = (m,\psi_i) \equiv \int_0^1 m(t)\psi_i(t)\, dt.$$

Proof: Let $\bar{P}_0$ and $\bar{P}_{10}$ be the completions of P_0 and P_{10} on $\bar{\mathfrak{B}}_{P_0}$ and $\bar{\mathfrak{B}}_{P_{10}}$ respectively. Then, from the definition, ξ_i, $i = 1, 2, \cdots$, are measurable with respect to both $\bar{\mathfrak{B}}_{P_0}$ and $\bar{\mathfrak{B}}_{P_{10}}$, and Gaussian distributed with respect to both $\bar{P}_0$ and $\bar{P}_{10}$ such that*

$$E_0\{\xi_i\} = E_{10}\{\xi_i - \nu_i\} = 0,$$

$$E_0\{\xi_i\xi_j\} = E_{10}\{(\xi_i - \nu_i)(\xi_j - \nu_j)\} = \lambda_i\delta_{ij}.$$

Furthermore, a modified version of Kauhunen-Loéve theorem† holds; namely, for every $t \in [0,1]$,

$$x_t - m_0(t) = \lim_{n\to\infty} \sum_{i=1}^{n} \xi_i\psi_i(t), \qquad \text{a.e. } (P_0). \tag{32}$$

Now, let $\mathfrak{B}_i'$, $i = 1, 2, \cdots$, be the minimal σ-field with respect to which ξ_i is measurable, and let

$$\mathfrak{B}' = \prod_{i=1}^{\infty} \mathfrak{B}_i'.$$

Then,

$$\mathfrak{B}_i' \subset \bar{\mathfrak{B}}_{P_0}, \mathfrak{B}_i' \subset \bar{\mathfrak{B}}_{P_{10}}, \mathfrak{B}' \subset \bar{\mathfrak{B}}_{P_0}, \mathfrak{B}' \subset \bar{\mathfrak{B}}_{P_{10}}. \tag{33}$$

Let $P_0'^{(i)}$ and $P_{10}'^{(i)}$ be the restrictions of $\bar{P}_0$ and $\bar{P}_{10}$ on $\mathfrak{B}_i'$, and P_0' and P_{10}' on $\mathfrak{B}'$. Then, it is readily seen that $P_0'^{(i)} \equiv P_{10}'^{(i)}$, $i = 1, 2, \cdots$, and

$$\frac{dP_{10}'^{(i)}}{dP_0'^{(i)}} = \exp\left[\frac{\xi_i^2}{2\lambda_i} - \frac{(\xi_i - \nu_i)^2}{2\lambda_i}\right] = \exp\left[\frac{\nu_i}{\lambda_i}\xi_i - \frac{\nu_i^2}{2\lambda_i}\right], \qquad \text{a.e. } (\bar{P}_0).$$

* E_0, E_{10} and E_1 denote expectations with respect to $\bar{P}_0$, $\bar{P}_{10}$ and $\bar{P}_1$ in general. However, if the function whose expectation is in question is $\mathfrak{B}$-measurable, the same symbols are used for expectations with respect to P_0, P_{10} and P_1 also.

† See Ref. 1, pp. 2801–2802.

Hence,

$$\int_\Omega \left(\frac{dP_{10}'^{(i)}}{dP_0'^{(i)}}\right)^{\frac{1}{2}} dP_0'^{(i)} = (2\pi\lambda_i)^{-\frac{1}{2}} \int_{-\infty}^{\infty} \exp\left(\frac{\nu_i}{2\lambda_i}\zeta - \frac{\nu_i^2}{4\lambda_i}\right) \exp\left(-\frac{\zeta^2}{2\lambda_i}\right) d\zeta$$

$$= \exp\left(-\frac{\nu_i^2}{8\lambda_i}\right).$$

Thus,

$$\prod_{i=1}^{\infty} \int_\Omega \left(\frac{dP_{10}'^{(i)}}{dP_0'^{(i)}}\right)^{\frac{1}{2}} dP_0'^{(i)} = \exp\left(-\frac{1}{8}\sum_{i=1}^{\infty}\frac{\nu_i^2}{\lambda_i}\right).$$

Hence, from Kakutani's theorem, either

$$P_0' \equiv P_{10}' \quad \text{or} \quad P_0' \perp P_{10}', \tag{34}$$

and $P_0' \equiv P_{10}'$ if and only if

$$\sum_{i=1}^{\infty}\frac{\nu_i^2}{\lambda_i} < \infty, \quad \text{i.e.,} \quad R_0^{-\frac{1}{2}} m \in \mathfrak{L}_2(0,1). \tag{35}$$

Next, for an arbitrary $t \in [0,1]$, define

$$\Gamma_t = \left\{\omega: x_t(\omega) - m_0(t) = \sum_{i=1}^{\infty} \xi_i(\omega)\psi_i(t)\right\},$$

$$\Lambda_t = \{\omega: x_t(\omega) - m_0(t) \in A\}, \qquad \Lambda_t' = \left\{\omega: \sum_{i=1}^{\infty} \xi_i(\omega)\psi_i(t) \in A\right\},$$

where A is an arbitrary Borel set. Put

$$\Lambda_t = (\Lambda_t \cap \Gamma_t) \cup (\Lambda_t \cap \Gamma_t^c), \qquad \Lambda_t' = (\Lambda_t' \cap \Gamma_t) \cup (\Lambda_t' \cap \Gamma_t^c).$$

Then, from (32)

$$\Lambda_t \cap \Gamma_t = \Lambda_t' \cap \Gamma_t, \qquad \bar{P}_0(\Lambda_t \cap \Gamma_t^c) = 0 = \bar{P}_0(\Lambda_t' \cap \Gamma_t^c).$$

Hence,

$$\bar{P}_0(\Lambda_t \triangle \Lambda_t') = 0.$$

Let $\mathfrak{B}'$ be a σ-field of sets of the form $\Lambda' \triangle N, K \in \mathfrak{B}', \bar{P}_0(N) = 0$. Then

$$\Lambda_t \in \bar{\mathfrak{B}}', 0 \leqq t \leqq 1.$$

That is, $x_t - m_0(t)$ is $\bar{\mathfrak{B}}'$-measurable for every $t \in [0,1]$. Hence, x_t is $\bar{\mathfrak{B}}'$-measurable for every t. But, since $\mathfrak{B}$ is the minimal σ-field with respect to which x_t is measurable for every t, we have

$$\mathfrak{B} \subset \bar{\mathfrak{B}}'. \tag{36}*$$

(ii) *Necessity:* Assume $P_0 \equiv P_{10}$. Then, $\bar{P}_0 \equiv \bar{P}_{10}$, thus $P_0' \equiv P_{10}'$.

* This part of the proof, i.e., establishment of (36), is not given in Ref. 4. In fact, Grenander's assertion is only on the primed measures P_0' and P_{10}'.

Hence, from (35),

$$R_0^{-\frac{1}{2}}m \in \mathfrak{L}_2(0,1).$$

Sufficiency: Assume $R_0^{-\frac{1}{2}}m \in \mathfrak{L}_2(0,1)$.

Then, from (35), $P_0' \equiv P_{10}'$. Then, from Lemma 2 (ii) together with (36),

$$P_0 \equiv P_{10}.$$

(i) *Dichotomy:* Assume P_0 and P_{10} are not equivalent. Then, from (ii), $R_0^{-\frac{1}{2}}m \notin \mathfrak{L}_2(0,1)$. Hence, from (35) and (34), $P_0' \perp P_{10}'$. Then, from Lemma 1 together with (33),

$$P_0 \perp P_{10}.$$

(iii) *Radon-Nikodym Derivative:* From (ii) and (35), $P_0 \equiv P_{10} \Rightarrow P_0' \equiv P_{10}'$. Then, from Lemma 2 (iv) together with (36), $dP_{10}/dP_0 = dP_{10}'/dP_0'$, a.e. $(\bar{P}_0)$. Then, the assertion follows from Lemma 3 (iii) with $\alpha_i = \beta_i = \lambda_i$, $i = 1, 2, \cdots$.

Corollary (Grenander):

If $R_0^{-1}m \in \mathfrak{L}_2(0,1)$, *then* $P_0 \equiv P_{10}$ *and*

$$\frac{dP_{10}}{dP_0} = \exp\left(x - \frac{m_0 + m_1}{2}, R_0^{-1}m\right), \qquad \textit{a.e.}\ (P_0).$$

Proof: The first assertion is obvious from Theorm 1 (ii) since

$$R_0^{-1}m \in \mathfrak{L}_2(0,1) \Rightarrow R_0^{-\frac{1}{2}}m \in \mathfrak{L}_2(0,1).$$

To prove the second assertion, note that

$$\sum_{i=1}^{n}\left(\xi_i - \frac{\nu_i}{2}\right)\frac{\nu_i}{\lambda_i} = \left(\sum_{i=1}^{n}\left(\xi_i - \frac{\nu_i}{2}\right)\psi_i, R_0^{-1}m\right).$$

Then, from (32) and the definition of ν_i, $i = 1, 2, \cdots$,

$$\left(x - \frac{m_0 + m_1}{2}, R_0^{-1}m\right) = \lim_{n\to\infty}\left(\sum_{i=1}^{n}\left(\xi_i - \frac{\nu_i}{2}\right)\psi_i, R_0^{-1}m\right),$$

a.e. $(\bar{P}_0)$.

APPENDIX C

Second Theorem on Equivalence and Singularity

Lemma 4: If either $R_1^{\frac{1}{2}}R_0^{-\frac{1}{2}}$ *or* $R_0^{\frac{1}{2}}R_1^{-\frac{1}{2}}$ *is unbounded, then* $P_0 \perp P_1$ *and* $P_{10} \perp P_1$.

Lemma 5: If $R_1^{\frac{1}{2}}R_0^{-\frac{1}{2}}$ *is bounded and* $R_0^{-\frac{1}{2}}m \in \mathfrak{L}_2(0,1)$, *then, for any sequence of functions* $f_i \in \mathfrak{L}_2(0,1)$, $i = 1, 2, \cdots$, *there exists a corresponding sequence of Gaussian variables* θ_i, $i = 1, 2, \cdots$, *(measurable with respect to* $\bar{\mathfrak{B}}_{P_0}$, $\bar{\mathfrak{B}}_{P_{10}}$ *and* $\bar{\mathfrak{B}}_{P_1}$*) such that for* $i, j = 1, 2, \cdots$,

$$\begin{aligned} E_0\{\theta_i + \nu_i\} &= E_{10}\{\theta_i\} = E_1\{\theta_i\} = 0, \\ E_0\{(\theta_i + \nu_i)(\theta_j + \nu_j)\} &= E_{10}\{\theta_i\theta_j\} = (f_i, f_j), \\ E_1\{\theta_i\theta_j\} &= (f_i, X^*Xf_j), \end{aligned} \tag{37}$$

where X *is the bounded extension of* $R_1^{\frac{1}{2}}R_0^{-\frac{1}{2}}$ *to the whole of* $\mathfrak{L}_2(0,1)$ *and* ν_i, $i = 1, 2, \cdots$, *are defined as*

$$\nu_i = (f_i, R_0^{-\frac{1}{2}}m).$$

Proof: Since $R_0^{\frac{1}{2}}(\mathfrak{L}_2(0,1))$ is dense in $\mathfrak{L}_2(0,1)$, there exists a sequence $\{f_{ij}\}_j$ for each f_i, $i = 1, 2, \cdots$, such that

$$R_0^{-\frac{1}{2}}f_{ij} \in \mathfrak{L}_2(0,1),\ j = 1, 2, \cdots, \text{ and } \lim_{j\to\infty} \| f_i - f_{ij} \| = 0,$$

where $\| f \|$ is the norm of f in the space $\mathfrak{L}_2(0,1)$. Then, through elementary steps, it can be shown that

$$\lim_{m,n\to\infty} (f_{im}, f_{jn}) = \lim_{n\to\infty}\lim_{m\to\infty} (f_{im}, f_{jn}) = (f_i, f_j). \tag{38}$$

1° Let θ_{ij}; $i,j = 1,2, \cdots$, be $\mathfrak{B}$-measurable functions such that

$$\theta_{ij} = (x - m_1, R_0^{-\frac{1}{2}}f_{ij}), \qquad \text{a.e. } (P_0, P_{10}, P_1).$$

Then, there exist random variables θ_i, $i = 1,2, \cdots$, which are measurable with respect to $\bar{\mathfrak{B}}_{P_0}$, $\bar{\mathfrak{B}}_{P_{10}}$ and $\bar{\mathfrak{B}}_{P_1}$, and Gaussian distributed with respect to $\bar{P}_0$, $\bar{P}_{10}$ and $\bar{P}_1$ such that

$$\theta_i = \text{l.i.m.}_{j\to\infty} \theta_{ij}, \qquad (\bar{P}_0, \bar{P}_{10}, \bar{P}_1).$$

To prove 1°, consider expectation with respect to P_0, P_{10} and P_1 of $|\theta_{ij} - \theta_{ik}|^2 = \theta_{ij}^2 - 2\theta_{ij}\theta_{ik} + \theta_{ik}^2$, $i = 1,2, \cdots$. First, note

$$\begin{aligned} E_0\{\theta_{ij}\theta_{ik}\} &= E_0\{(R_0^{-\frac{1}{2}}f_{ij}, x - m_0 - m)(x - m_0 - m, R_0^{-\frac{1}{2}}f_{ik})\} \\ &= (R_0^{-\frac{1}{2}}f_{ij}, R_0R_0^{-\frac{1}{2}}f_{ik}) - (R_0^{-\frac{1}{2}}f_{ij}, m)(R_0^{-\frac{1}{2}}f_{ik}, m). \end{aligned}$$

Thus, from (38),

$$\begin{aligned} \lim_{j,k\to\infty} E_0\{\theta_{ij}\theta_{ik}\} &= \lim_{j,k\to\infty} [(f_{ij}, f_{ik}) - (f_{ij}, R_0^{-\frac{1}{2}}m)(f_{ik}, R_0^{-\frac{1}{2}}m)] \\ &= \| f_i \|^2 - (f_i, R_0^{-\frac{1}{2}}m)^2. \end{aligned}$$

Hence,

$$\lim_{j,k\to\infty} E_0\{\,|\,\theta_{ij} - \theta_{ik}\,|^2\} = 0. \tag{39}$$

Secondly, note

$$\lim_{j,k\to\infty} E_{10}\{\theta_{ij}\theta_{ik}\} = \lim_{j,k\to\infty} (R_0^{-\frac{1}{2}} f_{ij}, R_0 R_0^{-\frac{1}{2}} f_{ik}) = \lim_{j,k\to\infty} (f_{ij}, f_{ik}) = \|f_i\|^2.$$

Hence,

$$\lim_{j,k\to\infty} E_{10}\{\,|\,\theta_{ij} - \theta_{ik}\,|^2\} = 0. \tag{40}$$

Thirdly, note

$$E_1\{\theta_{ij}\theta_{ik}\} = E_1\{(R_0^{-\frac{1}{2}} f_{ij}, R_1 R_0^{-\frac{1}{2}} f_{ik})\} = (Xf_{ij}, Xf_{ik}).$$

Since X is bounded, it is continuous. Thus, from (38),

$$\lim_{j,k\to\infty} (Xf_{ij}, Xf_{ik}) = \|Xf_i\|^2.$$

Hence,

$$\lim_{j,k\to\infty} E_1\{\,|\,\theta_{ij} - \theta_{ik}\,|^2\} = 0. \tag{41}$$

Next, upon combination of (39), (40) and (41), $\{\theta_{ij}\}_j$, $i = 1,2, \cdots$, are seen to be mean fundamental sequences with respect to $P_0 + P_{10} + P_1$. Hence, there exist θ_i, $i = 1,2, \cdots$, measurable with respect to $\overline{\mathfrak{B}}_{P_0+P_{10}+P_1}$ such that

$$\theta_i = \underset{j\to\infty}{\text{l.i.m.}}\ \theta_{ij}, \qquad (\bar{P}_0 + \bar{P}_{10} + \bar{P}_1).$$

But, since this implies

$$\theta_i = \underset{j\to\infty}{\text{l.i.m.}}\ \theta_{ij}, \qquad (\bar{P}_0, \bar{P}_{10}, \bar{P}_1), \qquad i = 1,2, \cdots,$$

θ_i, $i = 1,2, \cdots$, are measurable with respect to $\overline{\mathfrak{B}}_{P_0}$, $\overline{\mathfrak{B}}_{P_{10}}$ and $\overline{\mathfrak{B}}_{P_1}$, and are Gaussian distributed with respect to $\bar{P}_0$, $\bar{P}_{10}$ and $\bar{P}_1$.

2° To prove (37), simply note

$$\begin{aligned} E_0\{\theta_i + \nu_i\} &= \lim_{j\to\infty} E_0\{\theta_{ij}\} + \nu_i \\ &= \lim_{j\to\infty} E_0\{(x - m_0 - m, R_0^{-\frac{1}{2}} f_{ij})\} + \nu_i = 0, \\ E_{10}\{\theta_i\} &= \lim_{j\to\infty} E_{10}\{\theta_{ij}\} = 0, \end{aligned}$$

$$
\begin{aligned}
E_1\{\theta_i\} &= \lim_{j\to\infty} E_1\{\theta_{ij}\} = 0, \\
E_0\{(\theta_i + \nu_i)(\theta_j + \nu_j)\} &= \lim_{n\to\infty}\lim_{m\to\infty} E_0\{(\theta_{im} + \nu_i)(\theta_{jn} + \nu_j)\} \\
&= \lim_{n\to\infty}\lim_{m\to\infty} (R_0^{-\frac{1}{2}} f_{im}, R_0 R_0^{-\frac{1}{2}} f_{jn}) \\
&= (f_i, f_j), \\
E_{10}\{\theta_i\theta_j\} &= \lim_{n\to\infty}\lim_{m\to\infty} E_{10}\{\theta_{im}\theta_{jn}\} \\
&= \lim_{n\to\infty}\lim_{m\to\infty} (R_0^{-\frac{1}{2}} f_{im}, R_0 R_0^{-\frac{1}{2}} f_{jn}) \\
&= (f_i, f_j) \\
E_1\{\theta_i\theta_j\} &= \lim_{n\to\infty}\lim_{m\to\infty} E_1\{\theta_{im}\theta_{jn}\} \\
&= \lim_{n\to\infty}\lim_{m\to\infty} (R_0^{-\frac{1}{2}} f_{im}, R_1 R_0^{-\frac{1}{2}} f_{jn}) \\
&= (Xf_i, Xf_j),
\end{aligned}
$$

where (38) is used for the last three calculations.

Remark 1: The assertion of Lemma 5 with respect to P_{10} and P_1 only, is valid without the condition $R_0^{-\frac{1}{2}} m \in \mathfrak{L}_2(0,1)$.

Remark 2: Suppose $R_0^{-\frac{1}{2}} m \notin \mathfrak{L}_2(0,1)$ but there exist a sequence $\{f_{ij}\}_j$ for each f_i, $i = 1,2, \cdots$, such that $R_0^{-\frac{1}{2}} f_{ij} \in \mathfrak{L}_2(0,1)$, $j = 1,2, \cdots$,

$$\lim_{j\to\infty} \| f_i - f_{ij} \| = 0, \lim_{j\to\infty} (m, R_0^{-\frac{1}{2}} f_{ij}) = \nu_i'$$

for some real number ν_i'. Then, Lemma 5 is still valid if ν_i is replaced by ν_i', $i = 1, 2, \cdots$.

Remark 3: Suppose $R_0^{-\frac{1}{2}} m \notin \mathfrak{L}_2(0,1)$ and there is no such sequence. Then,

$$P_0 \perp P_1 .$$

Proof: Let

$$\nu_{ij} = (m, R_0^{-\frac{1}{2}} f_{ij}); \qquad i,j = 1,2, \cdots,$$

where

$$\lim_{j\to\infty} \| f_i - f_{ij} \| = 0$$

for each $i = 1,2, \cdots$. Without loss of generality, we assume that

$$\lim_{j\to\infty} |\nu_{ij}| = \infty$$

for some i. Define for such i,

$$\tilde{\theta}_{ij} = \theta_{ij}/\nu_{ij}, \qquad j = 1,2, \cdots .$$

Then

$$E_0\{\tilde{\theta}_{ij} + 1\} = E_1\{\tilde{\theta}_{ij}\} = 0, \qquad j = 1,2, \cdots .$$

Put

$$\sigma_{ij}^{\,0} = E_0\{(\tilde{\theta}_{ij} + 1)^2\}, \qquad \sigma_{ij}^{\,1} = E_1\{(\tilde{\theta}_{ij})^2\}.$$

Then,

$$\lim_{j\to\infty} \sigma_{ij}^{\,k} = 0, \qquad k = 0,1.$$

Thus, there exists a subsequence $\{\sigma_{ij_n}^{\,k}\}$ such that

$$\sum_{n=1}^{\infty} \sigma_{ij_n}^{\,k} < \infty, \qquad k = 0,1.$$

But, from Tchebycheff inequality, we have for some ε, $0 < \varepsilon < \frac{1}{2}$,

$$P_0(\{\omega: |\tilde{\theta}_{ij_n}(\omega)| < \varepsilon\}) < P_0(\{\omega: |\tilde{\theta}_{ij_n}(\omega) + 1| \geqq \varepsilon\}) \leqq \frac{\sigma_{ij_n}^{\,0}}{\varepsilon^2},$$

$$P_1(\{\omega: |\tilde{\theta}_{ij_n}(\omega)| \geqq \varepsilon\}) \leqq \frac{\sigma_{ij_n}^{\,1}}{\varepsilon^2}.$$

Hence, by Borel-Cantalli lemma,

$$P_0(\liminf_n\{\omega: |\tilde{\theta}_{ij_n}(\omega)| < \varepsilon\}) \leqq P_0(\limsup_n\{\omega: |\tilde{\theta}_{ij_n}(\omega)| < \varepsilon\}) = 0,$$

$$P_1(\limsup_n\{\omega: |\tilde{\theta}_{ij_n}(\omega)| \geqq \varepsilon\}) = 0.$$

Hence, by noting that

$$\limsup_n \{\omega: |\tilde{\theta}_{ij_n}(\omega)| \geqq \varepsilon\} = \Omega - \liminf_n\{\omega: |\tilde{\theta}_{ij_n}(\omega)| < \varepsilon\},$$

we have

$$P_0 \perp P_1.$$

Lemma 6: If $I - R_0^{-\frac{1}{2}}R_1R_0^{-\frac{1}{2}}$ *is a densely defined, bounded, completely continuous operator on* $\mathfrak{L}_2(0,1)$, *then*

$$x_t - m_1(t) = \lim_{m,n\to\infty}\left[\sum_{i=1}^{m}(R_0^{\frac{1}{2}}\varphi_i)(t)\eta_i + \sum_{i=1}^{n}(R_0^{\frac{1}{2}}\tilde{\varphi}_i)(t)\tilde{\eta}_i\right], \qquad (\mu \times \tilde{P}_{10})$$

where φ_i, $i = 1,2, \cdots$, *are the orthonormal eigenfunctions corresponding to nonzero eigenvalues of* $I - R_0^{-\frac{1}{2}}R_1R_0^{-\frac{1}{2}}$ *and* $\tilde{\varphi}_i$, $i = 1,2, \cdots$, *are an orthonormal basis of the null space of* $I - R_0^{-\frac{1}{2}}R_1R_0^{-\frac{1}{2}}$,* *and* η_i *and* $\tilde{\eta}_i$ *are defined by*

$$\left.\begin{aligned} \eta_i &= \underset{j\to\infty}{\text{l.i.m.}}\ (x - m_1, R_0^{-\frac{1}{2}}\varphi_{ij}), \\ \tilde{\eta}_i &= \underset{j\to\infty}{\text{l.i.m.}}\ (x - m_1, R_0^{-\frac{1}{2}}\tilde{\varphi}_i), \end{aligned}\right\} \quad (P_0, P_{10}, P_1),$$

where $\varphi_{ij}, \tilde{\varphi}_{ij} \in \mathfrak{L}_2(0,1)$; $i,j = 1,2, \cdots$, *are chosen in such a way that* $R_0^{-\frac{1}{2}}\varphi_{ij}$, $R_0^{-\frac{1}{2}}\tilde{\varphi}_{ij} \in \mathfrak{L}_2(0,1)$ *and, for each* i,

$$\lim_{j\to\infty} \| \varphi_i - \varphi_{ij} \| = 0, \qquad \lim_{j\to\infty} \| \tilde{\varphi}_i - \tilde{\varphi}_{ij} \| = 0,$$

and finally μ *is Lebesgue measure on Borel field* $\mathfrak{A}$ *of the subsets of* [0,1].

Proof: Note that φ_i and $\tilde{\varphi}_i$, $i = 1,2, \cdots$, exist since $I - R_0^{-\frac{1}{2}}R_1R_0^{-\frac{1}{2}}$ is densely defined, bounded, self-adjoint and completely continuous.

Consider

$$I_{m,n} = E_{10}\left\{\int_0^1 \left| x_t - m_1(t) - \sum_{i=1}^{m} (R_0^{\frac{1}{2}}\varphi_i)(t)\eta_i - \sum_{i=1}^{n} (R_0^{\frac{1}{2}}\tilde{\varphi}_i)(t)\tilde{\eta}_i \right|^2 dt\right\}.$$

By expanding the bracket,

$$\begin{aligned} I_{m,n} = {} & \int_0^1 r_0(t,t)\, dt - 2\sum_{i=1}^{m} \int_0^1 (R_0^{\frac{1}{2}}\varphi_i)(t) E_{10}\{[x_t - m_1(t)]\eta_i\}\, dt \\ & - 2\sum_{i=1}^{n} \int_0^1 (R_0^{\frac{1}{2}}\tilde{\varphi}_i)(t) E_{10}\{[x_t - m_1(t)]\tilde{\eta}_i\}\, dt \\ & + 2\sum_{i=1}^{m}\sum_{j=1}^{n} E_{10}\{\eta_i\tilde{\eta}_j\}(R_0^{\frac{1}{2}}\varphi_i, R_0^{\frac{1}{2}}\tilde{\varphi}_j) \\ & + \sum_{i,j=1}^{m} E_{10}\{\eta_i\eta_j\}(R_0^{\frac{1}{2}}\varphi_i, R_0^{\frac{1}{2}}\varphi_j) + \sum_{i,j=1}^{n} E_{10}\{\tilde{\eta}_i\tilde{\eta}_j\}(R_0^{\frac{1}{2}}\tilde{\varphi}_i, R_0^{\frac{1}{2}}\tilde{\varphi}_j). \end{aligned}$$

Note

$$\begin{aligned} E_{10}\{[x_t - m_1(t)]\eta_i\} = {} & E_{10}\left\{\underset{j\to\infty}{\text{l.i.m.}} \int_0^1 (R_0^{-\frac{1}{2}}\varphi_{ij})(s)[x_s - m_1(s)] \right. \\ & \left. \cdot [x_t - m_1(t)]\, ds\right\} \end{aligned}$$

* If the null space is finite dimensional, then $\{\tilde{\varphi}_i\}$ can be incorporated into $\{\varphi_i\}$ and there is no need to treat $\{\tilde{\varphi}_i\}$ separately.

$$= \lim_{j\to\infty} (R_0 R_0^{-\frac{1}{2}} \varphi_{ij})(t)$$
$$= (R_0^{\frac{1}{2}} \varphi_i)(t);$$

similarly,

$$E_{10}\{[x_t - m_1(t)]\tilde{\eta}_i\} = (R_0^{\frac{1}{2}} \tilde{\varphi}_i)(t).$$

Also, from Remark 1 of Lemma 5,*

$$E_{10}\{\eta_i \tilde{\eta}_j\} = (\varphi_i, \tilde{\varphi}_j) = 0,$$
$$E_{10}\{\eta_i \eta_j\} = (\varphi_i, \varphi_j) = \delta_{ij},$$
$$E_{10}\{\tilde{\eta}_i \tilde{\eta}_j\} = (\tilde{\varphi}_i, \tilde{\varphi}_j) = \delta_{ij}.$$

Therefore,

$$I_{m,n} = \int_0^1 r_0(t,t)\, dt - \sum_{i=1}^{m} (\varphi_i, R_0\varphi_i) - \sum_{i=1}^{n} (\tilde{\varphi}_i, R_0\tilde{\varphi}_i).$$

Now,

$$\int_0^1 r_0(t,t)\, dt = \sum_{k=1}^{\infty} \lambda_k.$$

On the other hand,

$$\sum_{i=1}^{\infty} (\varphi_i, R_0\varphi_i) + \sum_{i=1}^{\infty} (\tilde{\varphi}_i, R_0\tilde{\varphi}_i) = \sum_{i=1}^{\infty}\sum_{k=1}^{\infty} (\varphi_i, \psi_k)(R_0\varphi_i, \psi_k) + \sum_{i=1}^{\infty}\sum_{k=1}^{\infty} (\tilde{\varphi}_i, \psi_k)(R_0\tilde{\varphi}_i, \psi_k)$$
$$= \sum_{i=1}^{\infty}\sum_{k=1}^{\infty} \lambda_k (\varphi_i, \psi_k)^2 + \sum_{i=1}^{\infty}\sum_{k=1}^{\infty} \lambda_k (\tilde{\varphi}_i, \psi_k)^2$$
$$= \sum_{k=1}^{\infty} \lambda_k \left[\sum_{i=1}^{\infty} (\varphi_i, \psi_k)^2 + \sum_{i=1}^{\infty} (\tilde{\varphi}_i, \psi_k)^2\right]$$
$$= \sum_{k=1}^{\infty} \lambda_k.$$

Hence,

$$\lim_{m,n\to\infty} I_{m,n} = 0.$$

* Note that, it $R_0^{-\frac{1}{2}} R_1 R_0^{-\frac{1}{2}}$ is densely defined and bounded, then $R_1^{\frac{1}{2}} R_0^{-\frac{1}{2}}$ is bounded.

Lemma 7: Under the hypothesis of Lemma 6,

$$\mathfrak{B} \subset \bar{\hat{\mathfrak{B}}}$$

where $\bar{\hat{\mathfrak{B}}}$ *is a* σ*-field of sets of the form* $\hat{\Lambda} \triangle N$, $\hat{\Lambda} \in \hat{\mathfrak{B}}$, $\bar{P}_{10}(N) = 0$, *and* $\hat{\mathfrak{B}}$ *is the minimal* σ*-field with respect which all* η_i *and* $\tilde{\eta}_i$, $i = 1,2, \cdots$, *are measurable.*

Proof: It suffices to prove that x_t is $\bar{\hat{\mathfrak{B}}}$-measurable for every $t \in [0,1]$, since $\mathfrak{B}$ is the minimal σ-field with respect to which x_t is measurable for every t.

1° x_t is $\bar{\hat{\mathfrak{B}}}$-measurable for almost every t (with respect to μ).

To prove 1°, define

$$s_n(t,\omega) = \sum_{i=1}^{n} [(R_0^{\frac{1}{2}}\varphi_i)(t)\eta_i(\omega) + (R_0^{\frac{1}{2}}\tilde{\varphi}_i)(t)\tilde{\eta}_i(\omega)].$$

Then, from Lemma 6, there exists a subsequence $\{s_{n_k}(t,\omega)\}$ which converges to $x_t - m_1(t)$, a.e. $(\mu \times \bar{P}_{10})$. Namely, if

$$D = \{(t,\omega): x_t(\omega) - m_1(t) \neq \lim_{k\to\infty} s_{n_k}(t,\omega)\},$$

then

$$D \in \mathfrak{A} \times \bar{\mathfrak{B}}_{P_{10}} \quad \text{and} \quad (\mu \times \bar{P}_{10})(D) = 0.$$

Hence, from Fubini's theorem,* for almost every t

$$P_{10}(D_t) = 0,$$

where D_t is the section of D determined by t. In other words, $s_{n_k}(t,\omega)$ converges to $x_t(\omega) - m_1(t)$, a.e. $(\bar{P}_{10})$, for almost every t. Then, since each $s_{n_k}(t,\omega)$, $k = 1,2, \cdots$, is $\hat{\mathfrak{B}}$-measurable for every t, an argument analogous to the one in the proof of Theorem 1 (p. 1642) shows that $\Lambda_t = \{\omega: x_t(\omega) - m_1(t) \in A\}$ is $\bar{\hat{\mathfrak{B}}}$-measurable for almost every t, where A is any Borel set. Namely, $x_t - m_1(t)$ and, hence, x_t are $\bar{\hat{\mathfrak{B}}}$-measurable for almost every t.

2° x_t is $\bar{\hat{\mathfrak{B}}}$-measurable for every t.

To prove 2°, let $T \in \mathfrak{A}$ be a set of t for which x_t is $\bar{\hat{\mathfrak{B}}}$-measurable. Then, $\mu(T) = 1$. Since r_0 is continuous on $[0,1] \times [0,1]$ and T is dense in $[0,1]$, there exists for every $t \in [0,1]$ a sequence $\{t_n\}$, $t_n \in T$, converging to t such that

$$\lim_{n\to\infty} E_{10}\{|x_t - m_1(t) - x_{t_n} + m_1(t_n)|^2\} = 0.$$

* See Ref. 10, p. 147.

Hence, there exists a subsequence $\{t_{n_k}\}$ such that

$$\lim_{k\to\infty} [x_{t_{n_k}} - m_1(t_{n_k})] = x_t - m_1(t), \qquad \text{a.e. } (P_{10}).$$

Then, since each $x_{t_{n_k}} - m_1(t_{n_k})$, $k = 1,2, \cdots$, is $\bar{\mathfrak{B}}$-measurable for every t, the same argument used above shows that Λ_t is $\bar{\mathfrak{B}}$-measurable for every t. Namely, $x_t - m_1(t)$ and, hence, x_t are $\bar{\mathfrak{B}}$-measurable for every $t \in [0,1]$.

Theorem 2 (Pitcher):

(i) Either $P_{10} \equiv P_1$ *or* $P_{10} \perp P_1$,

(ii) $P_{10} \equiv P_1$ *if and only if* $I - R_0^{-\frac{1}{2}} R_1 R_0^{-\frac{1}{2}}$ *is a densely defined, bounded, completely continuous, Hilbert-Schmidt operator on* $\mathfrak{L}_2(0,1)$,

(iii) if $P_{10} \equiv P_1$,

$$\frac{dP_1}{dP_{10}} = \lim_{n\to\infty} \exp\left\{\frac{1}{2}\sum_{i=1}^{n}\left[\left(1 - \frac{1}{\rho_i}\right)\eta_i^2 - \log \rho_i\right]\right\}, \qquad \text{a.e. } (\bar{P}_{10}),$$

where ρ_i, $i = 1,2, \cdots$, *are the eigenvalues of* $R_0^{-\frac{1}{2}} R_1 R_0^{-\frac{1}{2}}$.

Proof:

(ii) *Necessity:* Assume $P_{10} \equiv P_1$.

Then, from Lemma 4, $R_1^{\frac{1}{2}} R_0^{-\frac{1}{2}}$ is bounded. Hence, $I - R_0^{-\frac{1}{2}} R_1 R_0^{-\frac{1}{2}}$ is densely defined and bounded.

The above statement implies that $R_0^{-\frac{1}{2}} R_1 R_0^{-\frac{1}{2}}$ is self-adjoint and positive-definite, and its bounded extension to the whole of $\mathfrak{L}_2(0,1)$ is equal to X^*X. Let $\int \nu dP_\nu$ be the spectral representation of X^*X. We now show by contradiction that X^*X has a purely discrete spectrum. Suppose for some $\varepsilon > 0$, $I - P_{1+\varepsilon}$ is infinite dimensional. Then, there exists a sequence $\{\nu_i\}$, $1 + \varepsilon \leqq \nu_1 < \nu_2 < \cdots$, and a sequence of orthonormal functions $f_i \in \mathfrak{L}_2(0,1)$, $i = 1, 2, \cdots$, such that

$$(P_{\nu_{i+1}} - P_{\nu_i}) f_i = f_i. \tag{42}$$

Hence, from Remark 1 of Lemma 5, there exists a sequence of random variables θ_i, $i = 1,2, \cdots$, which are measurable with respect to $\bar{\mathfrak{B}}_{P_{10}}$ and $\mathfrak{B}_{P_1}$ and Gaussian distributed with respect to both $\bar{P}_{10}$ and $\bar{P}_1$, such that

$$E_{10}\{\theta_i\} = E_1\{\theta_i\} = 0,$$

$$E_{10}\{\theta_i\theta_j\} = (f_i, f_j) = \delta_{ij},$$

$$E_1\{\theta_i\theta_j\} = (f_i, X^*Xf_j) = \delta_{ij}\int_{\nu_i}^{\nu_{i+1}} \nu d(f_i, P_\nu f_i) \geqq (1 + \varepsilon)\delta_{ij}.$$

Let $\mathfrak{B}^*$ be the minimal σ-field with respect to which all θ_i, $i = 1,2, \cdots$, are measurable, and let P_{10}^* and P_1^* be the restrictions of $\bar{P}_{10}$ and $\bar{P}_1$ on $\mathfrak{B}^*$. Then, from Lemma 3, $P_{10}^* \perp P_1^*$. It follows then from Lemma 1 that $P_{10} \perp P_1$, which is a contradiction. Therefore, $I - P_{1+\varepsilon}$ is finite dimensional for every $\varepsilon > 0$. Similarly, it can be shown that $P_{1-\varepsilon}$ is finite dimensional also. Hence, X^*X has a purely discrete spectrum, and 1 is the only limit point of the spectrum. Hence, $I - X^*X$ is completely continuous,* and so is $I - R_0^{-\frac{1}{2}}R_1R_0^{-\frac{1}{2}}$.

It follows from the preceding paragraph that the eigenvalues and the corresponding eigenfunctions, ρ_i and φ_i, $i = 1,2, \cdots$, of $R_0^{-\frac{1}{2}}R_1R_0^{-\frac{1}{2}}$ exist. Then, according to Lemma 5, η_i and $\tilde{\eta}_i$, $i = 1,2, \cdots$, defined in Lemma 6 have the following properties:

$$\begin{aligned}
E_{10}\{\eta_i\} &= E_{10}\{\tilde{\eta}_i\} = E_1\{\eta_i\} = E_1\{\tilde{\eta}_i\} = 0 \\
E_{10}\{\eta_i\eta_j\} &= (\varphi_i, \varphi_j) = \delta_{ij}, \\
E_{10}\{\tilde{\eta}_i\tilde{\eta}_j\} &= (\tilde{\varphi}_i, \tilde{\varphi}_j) = \delta_{ij}, \\
E_{10}\{\eta_i\tilde{\eta}_j\} &= (\varphi_i, \tilde{\varphi}_j) = 0, \\
E_1\{\eta_i\eta_j\} &= (\varphi_i, R_0^{-\frac{1}{2}}R_1R_0^{-\frac{1}{2}}\varphi_j) = \rho_i\delta_{ij}, \\
E_1\{\tilde{\eta}_i\tilde{\eta}_j\} &= (\tilde{\varphi}_i, R_0^{-\frac{1}{2}}R_1R_0^{-\frac{1}{2}}\tilde{\varphi}_j) = \delta_{ij} \\
E_1\{\eta_i\tilde{\eta}_j\} &= (\varphi_i, R_0^{-\frac{1}{2}}R_1R_0^{-\frac{1}{2}}\tilde{\varphi}_j) = (\varphi_i, \tilde{\varphi}_j) = 0.
\end{aligned} \tag{43}$$

Let $\hat{P}_{10}$ and $\hat{P}_1$ be the restrictions of $\bar{P}_{10}$ and $\bar{P}_1$ on $\mathfrak{B}$. Then, since $\rho_i > 0$, $i = 1,2, \cdots$, it follows from Lemma 3 that either $\hat{P}_{10} \equiv \hat{P}_1$ or $\hat{P}_{10} \perp \hat{P}_1$, and $\hat{P}_{10} \equiv \hat{P}_1$ if and only if

$$\sum_{i=1}^{\infty} \left(1 - \frac{1}{\rho_i}\right)^2 < \infty. \tag{44}$$

Furthermore, from Lemma 1, $\hat{P}_{10} \perp \hat{P}_1 \Rightarrow P_{10} \perp P_1$. But, since $P_{10} \equiv P_1$ from the hypothesis, we must have $\hat{P}_{10} \equiv \hat{P}_1$. Hence, (44) is satisfied, or equivalently,

$$\sum_{i=1}^{\infty} (1 - \rho_i)^2 < \infty. \tag{45}$$

Namely, $I - R_0^{-\frac{1}{2}}R_1R_1^{-\frac{1}{2}}$ is of Hilbert-Schmidt type.

Sufficiency: Assume that $I - R_0^{-\frac{1}{2}}R_1R_0^{-\frac{1}{2}}$ is a densely defined, bounded, completely continuous, Hilbert-Schmidt operator on $\mathfrak{L}_2(0,1)$. Then, $R_1^{\frac{1}{2}}R_0^{-\frac{1}{2}}$ is bounded, and $R_0^{-\frac{1}{2}}R_1R_0^{-\frac{1}{2}}$ is self-adjoint and positive-

*See Ref. 12, pp. 234–235.

definite. Thus, we establish η_i and $\tilde{\eta}_i$, $i = 1,2, \cdots$, and (43) as previously done. Now, since $I - R_0^{-\frac{1}{2}}R_1R_0^{-\frac{1}{2}}$ is of Hilbert-Schmidt type, (45) is satisfied, and so is (44). Then, since $\rho_i > 0$, $i = 1,2, \cdots$, it follows from Lemma 3 that $\hat{P}_{10} \equiv \hat{P}_1$. Then, from Lemma 7 and Lemma 2 (ii),

$$P_{10} \equiv P_1 .$$

(i) *Dichotomy:* Assume that P_{10} and P_1 are not equivalent. Then, one of the following three cases must hold:

(a) $I - R_0^{-\frac{1}{2}}R_1R_0^{-\frac{1}{2}}$ is either not densely defined or unbounded, or both,

(b) it is densely defined and bounded, but not completely continuous,

(c) it is densely defined, bounded and completely continuous, but not of Hilbert-Schmidt type.

In case (a), $R_1^{\frac{1}{2}}R_0^{-\frac{1}{2}}$ is unbounded. Hence, from Lemma 4, $P_{10} \perp P_1$. In case (b), X^*X has a spectral representation, and either $I - P_{1+\varepsilon}$ or $P_{1-\varepsilon}$ must be infinite dimensional for some $\varepsilon > 0$. Then, $P_{10}^* \perp P_1^*$ and, hence, $P_{10} \perp P_1$, as shown in the necessity part of the proof of (ii). In case (c), $I - R_0^{-\frac{1}{2}}R_1R_0^{-\frac{1}{2}}$ has the eigenvalues and eigenfunctions $1 - \rho_i$ and φ_i, $i = 1,2, \cdots$, and there are the associated Gaussian variables η_i and $\tilde{\eta}_i$, $i = 1,2, \cdots$, as described previously. But since $I - R_0^{-\frac{1}{2}}R_1R_0^{-\frac{1}{2}}$ is not of Hilbert-Schmidt type, (45) and, hence, (44) do not hold. Then, according to Lemma 3, $\hat{P}_{10} \perp \hat{P}_1$. Then, from Lemma 1, $P_{10} \perp P_1$. Therefore, we conclude that if P_{10} and P_1 are not equivalent then they must be singular.*

(ii) *Radon-Nikodym Derivative:* The assertion (iii) is an immediate consequence of Lemma 3 (iii) with $\nu_i = 0$, $i = 1,2, \cdots$, and Lemma 2 (*iv*).

Corollary 1: If $P_{10} \equiv P_1$ *and*

$$\left| \sum_{i=1}^{\infty} (1 - \rho_i) \right| < \infty ,$$

then

$$\frac{dP_1}{dP_{10}} = \left(\prod_{i=1}^{\infty} \rho_i \right)^{-\frac{1}{2}} \exp \left[\frac{1}{2} \sum_{i=1}^{\infty} \left(1 - \frac{1}{\rho_i} \right) \eta_i^2 \right], \qquad \text{a.e. } (\bar{P}_{10}).$$

Proof: Note that η_i, $i = 1,2, \cdots$, are mutually independent Gaussian variables with

$$E_{10}\{\eta_i\} = 0, \qquad E_{10}\{\eta_i^2\} = 1, \qquad E_{10}\{\eta_i^4\} = 3.$$

* Note this trivially implies that if P_{10} and P_1 are not singular, then they must be equivalent.

Hence

$$\left|\sum_{i=1}^{\infty} E_{10}\left\{\left(1-\frac{1}{\rho_i}\right)\eta_i^2\right\}\right| = \left|\sum_{i=1}^{\infty}\left(1-\frac{1}{\rho_i}\right)\right| < \infty,$$

$$\sum_{i=1}^{\infty} E_{10}\left\{\left(1-\frac{1}{\rho_i}\right)^2\eta_i^4\right\} = 3\sum_{i=1}^{\infty}\left(1-\frac{1}{\rho_i}\right)^2 < \infty.$$

Therefore,*

$$\sum_{i=1}^{\infty}\left(1-\frac{1}{\rho_i}\right)\eta_i^2 < \infty, \qquad \text{a.e. } (\bar{P}_{10}).$$

Then, the assertion follows upon combination of the above and Theorem 2 (iii).

Corollary 2 (Pitcher): If there exists a bounded, self-adjoint operator H on $\mathfrak{L}_2(0,1)$ satisfying

$$R_0HR_1 = R_1HR_0 = R_1 - R_0, \tag{46}$$

then

$$P_{10} \equiv P_1$$

and

$$\frac{dP_1}{dP_{10}} = \left(\prod_{i=1}^{\infty}\rho_i\right)^{-\frac{1}{2}} \exp\left[\tfrac{1}{2}(x-m_1, H(x-m_1)\right], \qquad \text{a.e. } (P_{10}).$$

APPENDIX D

Third Theorem on Equivalence and Singularity

Lemma 8: $P_{10} \perp P_1 \Rightarrow P_0 \perp P_1$.

Proof: If $P_{10} \perp P_1$, it follows from Theorem 2, (i) and (ii), that one of the three cases (a), (b) and (c) listed in the proof of Theorem 2 (i) holds.

In case (a), $R_1^{\frac{1}{2}}R_0^{-\frac{1}{2}}$ is unbounded, then $P_0 \perp P_1$ according to Lemma 4.

In case (b), at least, either $I - P_{1+\varepsilon}$ or $P_{1-\varepsilon}$ must be infinite dimensional for some $\varepsilon > 0$, as shown in the proof of Theorem 2 (ii). Suppose $I - P_{1+\varepsilon}$ is infinite dimensional. Then, there exists a sequence of orthonormal functions f_i, $i = 1,2, \cdots$, satisfying (42). Hence, according to Lemma 5, there exist a corresponding sequence of Gaussian variables

* See Ref. 5, p. 108.

θ_i, $i = 1, 2, \cdots$, such that

$$E_0\{\theta_i + \nu_i'\} = E_1\{\theta_i\} = 0, \qquad E_0\{(\theta_i + \nu_i')(\theta_j + \nu_j')\} = \delta_{ij},$$
$$E_1\{\theta_i\theta_j\} = (f_i, X^*Xf_j) \geqq (1 + \varepsilon)\delta_{ij},$$

provided that either $R_0^{-\frac{1}{2}}m \in \mathfrak{L}_2(0,1)$ or there exists a sequence $\{f_{ij}\}_j$, $i = 1,2, \cdots$, satisfying the conditions of Remark 2 of Lemma 5. Now, let $\mathfrak{B}^*$ be the minimal σ-field with respect to which all θ_i, $i = 1,2, \cdots$, are measurable, and let P_0^* and P_1^* be the restrictions of $\bar{P}_0$ and $\bar{P}_1$ on $\mathfrak{B}^*$. Then, from Lemma 3 and the above result, it follows that $P_0^* \perp P_1^*$. Hence, from Lemma 1, $P_0 \perp P_1$. On the other hand, suppose neither $R_0^{-\frac{1}{2}}m \in \mathfrak{L}_2(0,1)$ nor there exist such a sequence $\{f_{ij}\}_j$ for some i. Then, from Remark 3 of Lemma 5, $P_0 \perp P_1$ also.

Similarly, if $P_{1-\varepsilon}$ is infinite dimensional, it can be shown that $P_0 \perp P_1$.

In case (c), we can assume existence of the Gaussian variables η_i and $\tilde{\eta}_i$, $i = 1,2, \cdots$, with the properties (43) and the following:

$$E_0\{\eta_i + \gamma_i\} = E_0\{\tilde{\eta}_i + \tilde{\gamma}_i\} = 0,$$
$$E_0\{(\eta_i + \gamma_i)(\eta_j + \gamma_j)\} = E_0\{(\tilde{\eta}_i + \tilde{\gamma}_i)(\tilde{\eta}_j + \tilde{\gamma}_j)\} = \delta_{ij}, \quad (47)$$
$$E_0\{(\eta_i + \gamma_i)(\tilde{\eta}_j + \tilde{\gamma}_j)\} = 0,$$

where $\gamma_i = (\varphi_i, R_0^{-\frac{1}{2}}m)$, $\tilde{\gamma}_i = (\tilde{\varphi}_i, R_0^{-\frac{1}{2}}m)$, $i = 1, 2, \cdots$.

Since $I - R_0^{-\frac{1}{2}}R_1R_0^{-\frac{1}{2}}$ is not of Hilbert-Schmidt type, (45) does not hold. Thus, (44) is not satisfied. Hence, from Lemma 3, $\hat{P}_0 \perp \hat{P}_1$. Then, from Lemma 1, $P_0 \perp P_1$.

Lemma 9: $P_0 \perp P_{10} \Rightarrow P_0 \perp P_1$.

Proof: Since $P_0 \perp P_{10}$, there exist a non-empty set $\Lambda \in \mathfrak{B}$ such that

$$P_0(\Omega - \Lambda) = 0 \quad \text{and} \quad P_{10}(\Lambda) = 0.$$

Now, if $P_{10} \equiv P_1$, then $P_1(\Lambda) = 0$. Hence, we have

$$P_0(\Omega - \Lambda) = 0 \quad \text{and} \quad P_1(\Lambda) = 0,$$

namely, $P_0 \perp P_1$. If P_{10} and P_1 are not equivalent, then they must be singular according to Theorem 2 (i), i.e., $P_{10} \perp P_1$. Then, from Lemma 8, $P_0 \perp P_1$.

Theorem 3:

(*i*) *Either* $P_0 \equiv P_1$ *or* $P_0 \perp P_1$,
(*ii*) $P_0 \equiv P_1$ *if and only if*

(a) $I - R_0^{-\frac{1}{2}} R_1 R_0^{-\frac{1}{2}}$ *is a densely defined, bounded, completely continuous, Hilbert-Schmidt operator on* $\mathfrak{L}_2(0,1)$,

(b) $R_0^{-\frac{1}{2}} m \in \mathfrak{L}_2(0,1)$,

(*iii*) *if* $P_0 \equiv P_1$,

$$\frac{dP_1}{dP_0} = \exp\left\{\frac{1}{2}\sum_{i=1}^{\infty}\left[\left(1 - \frac{1}{\rho_i}\right)\eta_i^2 - \log \rho_i\right]\right\} \cdot \exp\left\{\sum_{i=1}^{\infty}\left[\gamma_i\left(\eta_i + \frac{\gamma_i}{2}\right) + \tilde{\gamma}_i\left(\tilde{\eta}_i + \frac{\tilde{\gamma}_i}{2}\right)\right]\right\}, \qquad \text{a.e. } (\tilde{P}_0).$$

(Remark) Note it follows from Theorems 1 and 2 that the necessary and sufficient condition for $P_0 \equiv P_1$ is (a) $P_{10} \equiv P_1$ and (b) $P_0 \equiv P_{10}$.

Proof:

(ii) *Necessity:* Assume $P_0 \equiv P_1$.

Then, from Theorem 2 (i) and Lemma 8,

$$P_{10} \equiv P_1,$$

while, from Theorem 1 (i) and Lemma 9,

$$P_0 \equiv P_{10}.$$

Hence, (a) and (b) follow immediately from Theorem 2 (ii) and Theorem 1 (ii) respectively.

Sufficiency: Obvious since $P_0 \equiv P_{10}$ and $P_{10} \equiv P_1$ imply $P_0 \equiv P_1$.

(i) *Dichotomy:* Assume that P_0 and P_1 are not equivalent.

Then, it follows from the sufficiency part of (ii) as well as from Theorem 2 (i) and Theorem 1 (i) that, at least, either

$$P_{10} \perp P_1 \quad \text{or} \quad P_0 \perp P_{10}.$$

Then, from Lemma 8 and Lemma 9, we have

$$P_0 \perp P_1.$$

Thus, if P_0 and P_1 are not equivalent, then they must be singular.

(iii) *Radon-Nikodym Derivative:* From Lemma 3 (iii) and Lemma 2 (iv), in conjunction with (43) and (47), we have

$$\frac{dP_1}{dP_0} = \exp\left\{\sum_{i=1}^{\infty}\left[\frac{1}{2}\left(1 - \frac{1}{\rho_i}\right)\eta_i^2 - \tfrac{1}{2}\log \rho_i + \gamma_i\left(\eta_i + \frac{\gamma_i}{2}\right) + \tilde{\gamma}_i\left(\tilde{\eta}_i + \frac{\tilde{\gamma}_i}{2}\right)\right]\right\}, \qquad \text{a.e. } (\tilde{P}_0). \tag{48}$$

Since $P_0 \equiv P_1 \Rightarrow P_0 \equiv P_{10}$ and $P_{10} \equiv P_1$ according to (ii), it follows from Theorem 2 (iii) that

$$\sum_{i=1}^{\infty} \left[\left(1 - \frac{1}{\rho_i}\right) \eta_i^2 - \log \rho_i \right] < \infty, \qquad \text{a.e. } (\bar{P}_0).$$

Hence, the remainder of the exponent of (48) converges a.e. $(\bar{P}_0)$. This proves (iii).

Corollary 1: If $P_0 \equiv P_1$ *and*

$$\left| \sum_{i=1}^{\infty} (1 - \rho_i) \right| < \infty,$$

then

$$\frac{dP_1}{dP_0} = \left(\prod_{i=1}^{\infty} \rho_i \right)^{-\frac{1}{2}} \exp\left[\frac{1}{2} \sum_{i=1}^{\infty} \left(1 - \frac{1}{\rho_i}\right) \eta_i^2 \right] \cdot \exp\left\{ \sum_{i=1}^{\infty} \left[\gamma_i \left(\eta_i + \frac{\gamma_i}{2} \right) + \tilde{\gamma}_i \left(\tilde{\eta}_i + \frac{\tilde{\gamma}_i}{2} \right) \right] \right\}, \qquad \text{a.e. } (\bar{P}_0).$$

Proof: This follows from Corollary 1 of Theorem 2 and Theorem 3 (iii).

Corollary 2: If there exists a bounded, self-adjoint operator H *on* $\mathfrak{L}_2(0,1)$ *satisfying* (46), *and* $R_0^{-\frac{1}{2}} m \in \mathfrak{L}_2(0,1)$, *then*

(i) $P_0 \equiv P_1$,

(ii) $$\frac{dP_1}{dP_0} = \left(\prod_{i=1}^{\infty} \rho_i \right)^{-\frac{1}{2}} \exp\left[\tfrac{1}{2}(x - m_1, H(x - m_1)) + \left(x - \frac{m_0 + m_1}{2}, R_0^{-1} m \right) \right], \qquad \text{a.e. } (P_0),$$

(iii) $R_0(\mathfrak{L}_2(0,1)) = R_1(\mathfrak{L}_2(0,1))$.

Proof:

(i) The assertion is an immediate consequence of combination of Theorem 3 and Corollary 2 of Theorem 2.

(ii) Note

$$\frac{dP_1}{dP_0} = \frac{dP_1}{dP_{10}} \frac{dP_{10}}{dP_0}, \qquad \text{a.e. } (P_0).$$

Then, the assertion follows upon combination of Corollary 2 of Theorem 2 and the corollary to Theorem 1.

(iii) From (46),

$$R_1(\mathfrak{L}_2(0,1)) = [R_0(HR_1 + I)](\mathfrak{L}_2(0,1)) \subset R_0(\mathfrak{L}_2(0,1)),$$

$$R_0(\mathfrak{L}_2(0,1)) = [R_1(I - HR_0)](\mathfrak{L}_2(0,1)) \subset R_1(\mathfrak{L}_2(0,1)).$$

Hence, the assertion follows.

REFERENCES

1. Kadota, T. T., Optimum Reception of Binary Gaussian Signals, B.S.T.J., *43*, Nov., 1964, pp. 2767–2810.
2. Rao, C. R., Varadarajan, V. S., Discrimination of Gaussian Processes, Sankhyā, Ser. A, *25*, 1963, pp. 303–330.
3. Pitcher, T. S., An Integral Expression for The Log Likelihood Ratio of Two Gaussian Processes, to appear in J. SIAM.
4. Grenander, U., Stochastic Processes and Statistical Inference, Arkiv für Matematik, *17*, 1950, pp. 195–277.
5. Doob, J. L., *Stochastic Processes*, John Wiley & Sons, New York, 1953.
6. Hajek, J., A Property of J-Divergence of Marginal Probability Distributions, Czechoslovak Math. J., *83*, 8, 1958, pp. 460–463.
   Hajek, J., On a Property of Normal Distribution of any Stochastic Process, (in Russian), Czechoslovak Math. J., *83*, 8, 1958, pp. 610–618. (A translation appears in Selected Translations in Mathematical Statistics and Probability, Institute of Mathematical Statistics and American Mathematical Society, *1*, pp. 245–252).
7. Feldman, J., Equivalence and Perpendicularity of Gaussian Processes, Pacific J. Math., *8*, No. 4, 1958, pp. 699–708.
   Feldman, J., Correction to Equivalence and Perpendicularity of Gaussian Processes, Pacific J. Math., *9*, No. 4, 1959, pp. 1295–1296.
8. Shepp, L. A., Gaussian Measures in Function Space, (to appear in Pacific J. Math).
9. Root, W. L., *Singular Gaussian Measures in Detection Theory*, Proc. of Symposium on Time Series Analysis, John Wiley, New York, 1963, pp. 292–315.
10. Halmos, P. R., *Measure Theory*, Van Nostrand, Princeton, 1950.
11. Apostol, T. M., *Mathematical Analysis*, Addison-Wesley, Reading, Mass., 1957.
12. Riess, F., and Sz-Nagy, B., *Functional Analysis*, Frederick Ungar Publishing Co., New York, 1955.

0.63μ Scatter Measurements from Teflon* and Various Metallic Surfaces

By R. A. SEMPLAK

(Manuscript received May 27, 1965)

Angular scatter measurements obtained by illuminating Teflon and various metallic surfaces at normal incidence with a 0.63μ laser beam are discussed. A method for measuring the scatter in the specular direction is also presented. The measured data are found to be in good agreement with models, which take into account specular reflection, scatter, and absorption. The surfaces investigated were aluminum, steel, brass, first surface aluminized mirror, and aluminized roughened glass surfaces.

I. INTRODUCTION

A cursory examination of a visible laser beam incident on any refracting or reflecting system will show that energy is scattered from the laser beam. This scattering may be due to dust particles, irregularities of the surfaces, or to inhomogeneities in the volume of the material. Depending upon the surface, some energy is scattered at wide angles from the laser beam. A knowledge of the amount of scattered energy is of primary importance in determining the losses and may assist in characterizing the surface.

II. EQUIPMENT

To measure the power lost by scattering and reflection, a probing type detector is desirable. The detecting system must in no way affect the beam propagation and must be able to discriminate and measure the energy flowing across a small but finite area. An ideal probe-detector is visualized here as a single mode optical amplifier and lens system that would respond only to plane waves within a diffraction limited spot in the focal plane of the lens system.

In view of the nonavailability of an ideal probe, an effort has been made to design and fabricate a detecting system which would embody

* Registered trademark of E. I. DuPont de Nemours, Inc.

as many features of the ideal case as possible. For lack of a better name this system will be referred to as a focused optical probe.

The basic elements of the focused probe and its general configuration are shown in Fig. 1(a). By using a lens and mirror system as an adjunct to the detector (an RCA 7102 photomultiplier), well defined measurements can be obtained. As indicated in Fig. 1(a), a mirror has been used to reduce the overall length of the probe and to permit scatter measurements to be made to within a few degrees of an incident beam. An iris centered in the focal plane of the lens nearest the detector acts as a stop which limits the acceptance angle of that lens and provides a small area at the focal point of the second lens for measure-

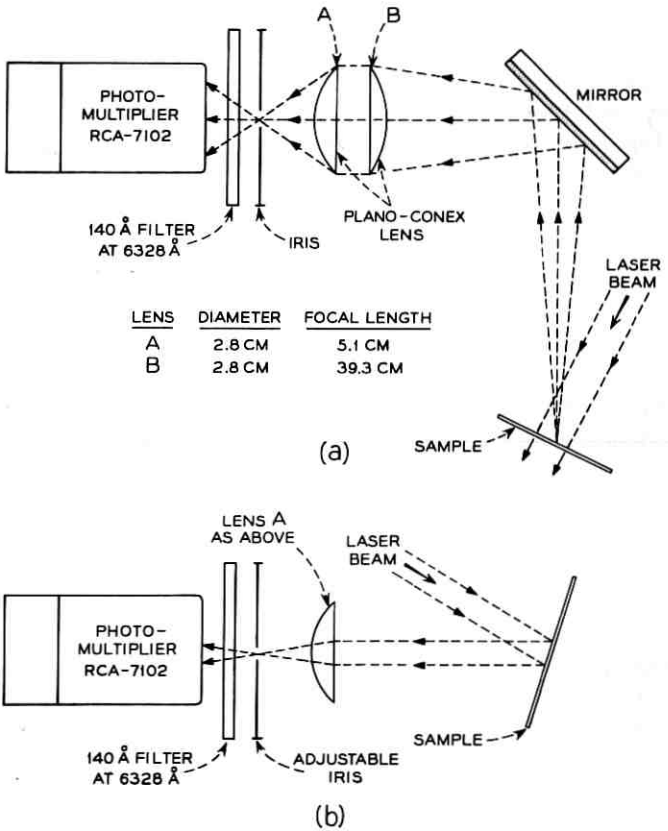


Fig. 1 — Optical probes; (a) focused, for angular scattering measurements, (b) modified for measurement of scattering in the direction of specular reflection.

ment of the power reflected or scattered from a surface placed in that focal plane.

The elements of the probe are assembled in a light-tight housing mounted on a cross-feed indexing table which permits one rotary and two transverse motions for probe positioning. The position of the focal point is indicated by a gauge attached to the mirror housing. Also shown here as Fig. 1(b), is the drawing of the probe modified for measurement of scattering in the specular direction; this modified version is discussed later.

To provide an adequate measuring range in the detection system, a chopper and phase detector are used, as indicated in Fig. 2. With the probe iris set at minimum opening the signal to noise ratio for this equipment is about 75 db using a one-sec. time constant. The dc excited laser (length — 1 meter) is operated at 0.63μ; its cavity mirrors have a radius of curvature of 10 meters and an iris within the cavity is used to suppress higher order modes. Wratten neutral density filters are used at attenuators.

The focused probe response was measured by rotating the focal area of the probe about a fixed point in the laser beam, as indicated in Fig. 3. The measured response is also shown there. The response falls rapidly as θ, the angle between the laser beam and the normal to the focal area of the probe, increases. Also plotted as a dashed curve in

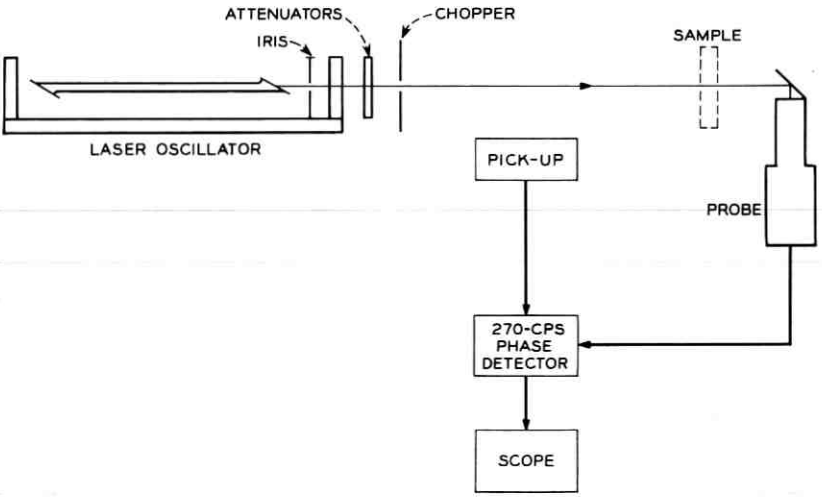


Fig. 2 — Equipment schematic.

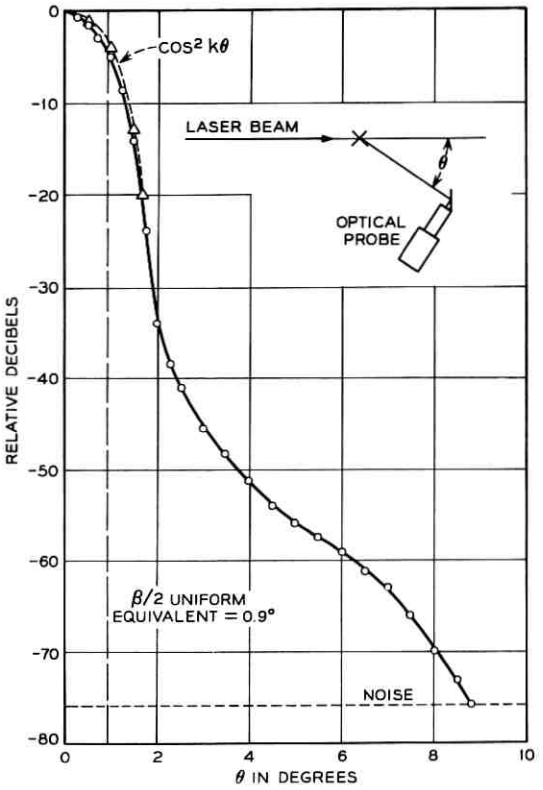


Fig. 3 — Focused probe response curve.

Fig. 3 is the function $\cos^2 k\theta$ which is a reasonable approximation of the measured response (if $k = 50$). Later discussion will be concerned with the idealized case of a uniform distribution over the solid angle of the probe. If the actual response, $\cos^2 k\theta$, is integrated over θ, one can determine the width of the equivalent uniform response; this turns out to be $\beta \cong 2 \times 0.9° = 1.8°$ as indicated in Fig. 3.

III. TEFLON

3.1 *Measurements*

Teflon was one of the first materials to be tested. The measurements indicate that scattering from Teflon is diffuse. They produce a Lambert

type pattern which can be represented reasonably well by a simple mathematical model obeying a cosine law.

The scattering measurements were made as follows: a particular thickness of Teflon was introduced normal to the laser beam (Fig. 2) and probe measurements were made by rotating the axis of the probe about a selected point on the surface of the sample. In the following, a reference to forward scatter means that the measurements were made with the probe situated in the forward hemisphere defined by the direction of propagation of the laser beam whereas back scatter means that these measurements were made with the probe located in the back hemisphere.

Teflon pieces of $\frac{1}{16}$, $\frac{1}{8}$, $\frac{1}{4}$, $\frac{1}{2}$, and of 1-inch thickness were used as samples. As one can see from Fig. 4, where the back-scatter data obtained from several of the samples are plotted (upper curve), any one of the samples can be well represented by the average curve shown there, hence this average curve will be used in what follows. It should be noted that the portion of the curve from $\theta = 175°$ to $\theta = 180°$ was obtained by extrapolation. Also shown in Fig. 4 are the forward scatter data from the one-half inch sample and for comparison, the computed cosine dependence for a Lambert surface (the dashed curve). Since the samples were of finite thicknesses, measurements were not taken out to $\theta = 90°$.

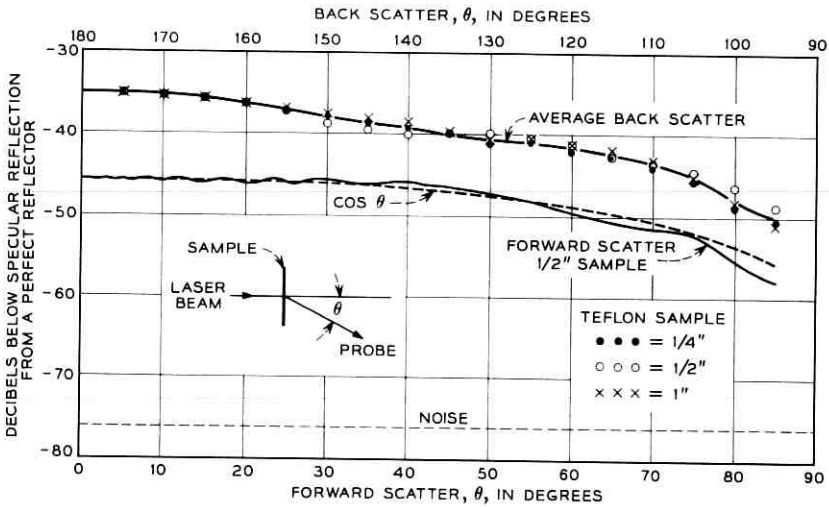


Fig. 4 — Forward and back scatter measurements from Teflon samples.

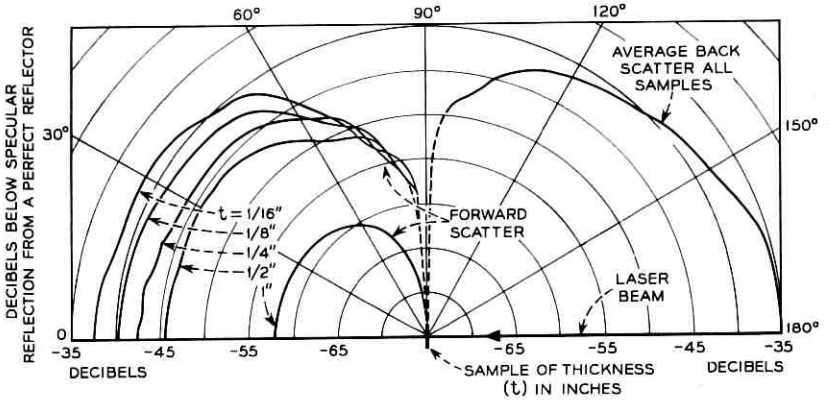


Fig. 5 — Polar plot of scattering from Teflon samples of various thicknesses (t).

Fig. 5 shows both forward and back scatter data for all sample thicknesses as polar plots for ease of comparison.

3.2 *Discussion of the Scatter Model for Teflon*

First, consider unit power incident on the surface of a partially transparent material and assume there is no specularly reflected component[1] from the surface; the energy is either scattered or absorbed. Let $\sigma(\Omega)$ be the scattering coefficient per unit solid angle, then by conservation of energy

$$1 = \frac{1}{4\pi} \int_{4\pi} \sigma(\Omega)\, d\Omega + a \tag{1}$$

where a is the fraction of the incident power that is absorbed in the sample and the integral term represents all of the scattered power, and in this case includes scattering from the volume of the material.

Now, assuming azimuthal symmetry in the scattered power and letting $\sigma(\theta)$ be the scattering coefficient in the direction $\Omega(\theta, \varphi)$, then the power $S_m(\theta)$ accepted by the probe (relative to that which would be received from a suitably oriented perfect reflector) looking to the direction θ^* is

$$S_m(\theta) = \sigma(\theta)\, \frac{\omega_0}{4\pi} \tag{2}$$

where $\omega_0 \cong \beta^2$ is the solid angle of the probe response.

Let $\sigma(\theta) = \sigma_b(\theta) + \sigma_f(\theta)$ where the first and second terms on the

* θ is the angle measured between the laser beam and the normal to the focal area of the probe (Fig. 4).

right represent the back and forward scatter coefficients respectively. Substituting for $\sigma(\theta)$ in (1),

$$1 = \frac{1}{2}\int_{\pi/2}^{\pi} \sigma_b(\theta) \sin\theta \, d\theta + \frac{1}{2}\int_{0}^{\pi/2} \sigma_f(\theta) \sin\theta \, d\theta + a. \qquad (3)$$

From (2), $\sigma(\theta) = 4\pi S_m(\theta)/\omega_0$ and since the measurements are made in both the forward and back scattering directions one writes $S_m(\theta)$ as $S_{mf}(\theta)$ or $S_{mb}(\theta)$. Substituting for σ in (3), one obtains

$$1 = \frac{2\pi}{\omega_0}\int_{\pi/2}^{\pi} S_{mb}(\theta) \sin\theta \, d\theta + a + \frac{2\pi}{\omega_0}\int_{0}^{\pi/2} S_{mf}(\theta) \sin\theta \, d\theta. \qquad (4)$$

From numerical integration of the data, it has become evident that for all Teflon samples, one-half of the incident power is scattered Lambert-wise in the *back hemisphere*, i.e., the first integral of (4) equals one-half. The third term of (4) is evaluated by numerically integrating the data. Subtracting these two terms from unity gives that fraction a of the power absorbed by the sample.

The values for the total power scattered in the forward direction (the third term in (4)) as obtained by numerical integration are tabulated in Table I.

3.3 *Evaluation of Absorption Coefficient*

The power absorbed by the sample is written as

$$a(db) = 10 \log_{10} \exp(-\alpha d) \qquad (5)$$

where α is an absorption coefficient and d the thickness of the sample. Since the total forward scattered power is known from measurements similar to those in Fig. 4, and assuming again that half of the power is available for transmission, the power absorbed by the sample can be determined by solving (5) for α. By using the values given in Table I and averaging the α's obtained from (5), this gives an absorption coefficient for Teflon of $\alpha = 4.7$ inches^{-1} (± 1.1).

TABLE I

Teflon SampleThickness (inches)	Total Forward Scatter
$\frac{1}{16}$	0.38
$\frac{1}{8}$	0.24
$\frac{1}{4}$	0.136
$\frac{1}{2}$	0.085
1	0.005

IV. METALLIC SURFACES

Samples of aluminum, steel, and brass were selected from stock of bulk metals, the only criterion applied in the selection being that the samples be reasonably flat. One side of each sample was cleaned and polished with a liquid metal cleaner. In addition to the bulk metals, a first surface aluminized mirror and three aluminized roughened glass flats were measured. The roughened glass flats were prepared by hand grinding each flat with one of the following grit sizes: 5μ, 12μ, or 25μ. After cleaning, the surfaces were aluminized in vacuo.

Using the focused probe (Fig. 1(a)), angular scattering measurements were made as follows: a particular sample was introduced normal to the laser beam and probe measurements were made by rotating the axis of the probe about a selected point on the surface of the sample, as shown on Fig. 6(a). The minimum angle θ (as measured between the beam and the axis of the probe) at which angular scattering could be measured was about 175°. The measured (angular) scatter data for the aluminum, steel, brass, first surface mirror, and the roughened glass flats are plotted in Figs. 6(a) through 6(d), respectively. The solid curves represent the power measured by the focused probe (normalized to the specularly reflected power the probe would have measured if the sample were replaced by a perfect reflector) for the samples just mentioned as the probe is rotated about the sampling point from $\theta = 175°$ to $\theta = 95°$. The solid dots at $\theta = 180°$ represent the measurements in the specular direction. It should be noted that this value (at $\theta = 180°$) contains both the specularly reflected component as well as the scattered power in the specular direction. The amount of energy scattered in the specular direction is obtained by a modification of the focused probe (Fig. 1(b)), which will be discussed later. With the aid of this second set of data, one can find the specular reflection coefficient for the surface. The data obtained using these two methods permit one to evaluate the total scattered energy.

In Figs. 6(a), 6(b), and 6(c), there are two sets of curves for each side of the metal samples. The curves labelled P and D (without subscripts) represent the polished and dull sides for a given orientation of the sample; the second set (designated by subscript 90) was obtained by rotating the sample 90° about its normal. From these figures, the effects of preferential scattering are readily evident. This type of scattering is due to orientations given to the surface facets as a result of the rolling operations used in processing the bulk metal.[2] As

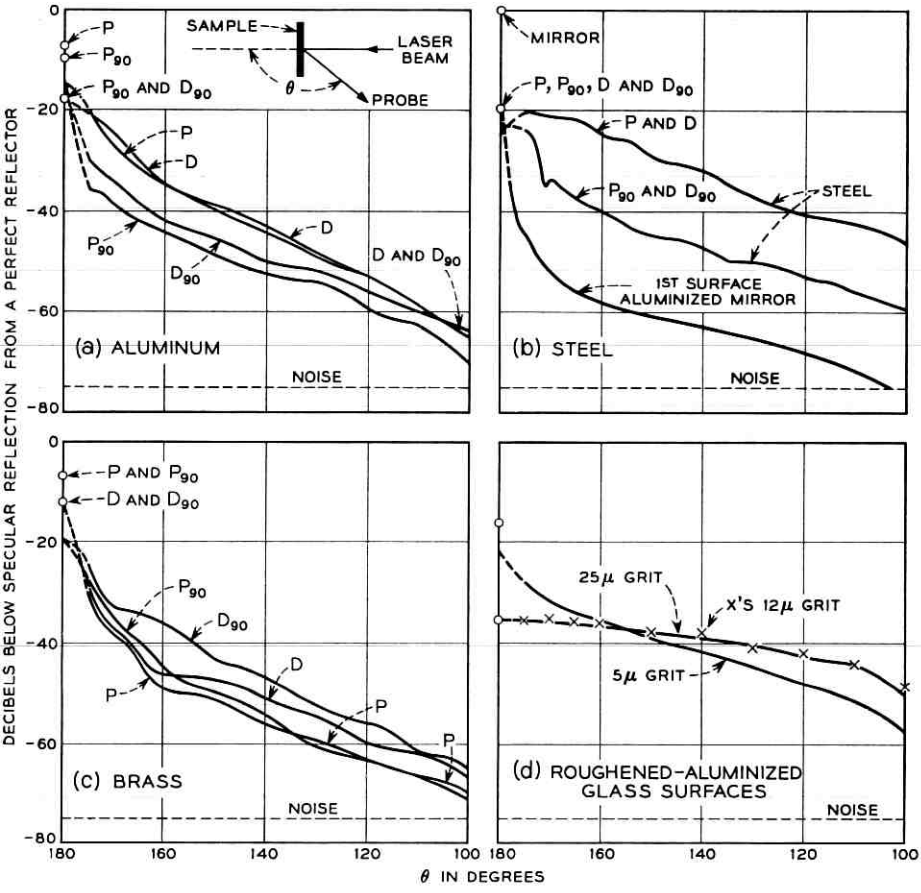


Fig. 6 — Angular scatter measurements for various surfaces. P and D denote the polished and dull side of sample. The subscript 90 denotes a 90° rotation of the sample about its normal. The solid dots at $\theta = 180°$ are total measured values, i.e., they include the specularly reflected component.

the number of rolling operations used to achieve the final metal thickness are increased, the more anisotropic the surface becomes. Also shown in Fig. 6(b), scattering from a first surface mirror is relatively small as expected. From Fig. 6(d), it appears that the 12μ and 25μ grit roughened flats are diffuse scatterers.

It becomes rather apparent after studying these figures that if the scattered energy could be measured in the specular direction, a good approximation of the total scattered energy could be obtained by

smoothly connecting the curve from $\theta = 175°$ to the value obtained at $\theta = 180°$ and then integrating, assuming circular symmetry in the scattered component. A method of measuring scattering in this direction ($\theta = 180°$) will be discussed next.

V. SCATTERING IN THE SPECULAR DIRECTION

If the assumption is made that the energy in the specular direction can be described by a specularly reflected component and an isotropically diffuse component, a method for measuring the scatter in the specular direction presents itself. For this purpose, a single plano-convex lens with a variable iris in its focal plane can be used. Since the specularly reflected component is focused, increase in iris opening* will not change the transmission through the systems and any increase in measured level will be due to the scattering in the specular direction. This increase will be proportional to the iris area provided the scattering in the specular direction is sufficiently diffuse. Operation of such a system can quickly be checked by viewing a plane wave (say the laser beam) directly in which case the energy measured will be constant, i.e., independent of iris openings, whereas for a perfectly diffuse surface, the energy measured would obey a square law in the diameter d of the iris.

The focused optical probe (Fig. 1(a)) discussed above has been modified to permit its use in measuring this scattering in the specular direction. As shown in Fig. 1(b), this modification is accomplished by merely removing the mirror and lens B. Measurements looking directly at the laser beam using the modified probe (Fig. 1(b)) show no appreciable change in level as the iris opening is varied from minimum to maximum. To check the square law performance of the system, scattering from a teflon surface was measured — since previous measurements (Section III) had shown it to be a very diffuse scatterer. The data obtained from the Teflon sample are shown as a solid curve in Fig. 7; these, when compared with the calculated (dashed) square-law curve, indicate that the system responds properly to a completely diffuse field.†

* Iris openings much larger than the diffraction limit of the lens are considered here.

† In some cases, it was necessary to use an incident beam larger in diameter than the normal laser beam. A lens combination was used to obtain the desired beam magnification. Measurements of the response of the modified probe to this enlarged beam show a slight increase of 0.3 db in measured power as the iris opening is increased from 0.02-inches diameter (minimum) to 0.14-inches diameter (maximum). This increase is attributed to scattering in the lenses used to enlarge the beam and has negligible effects on the measurements.

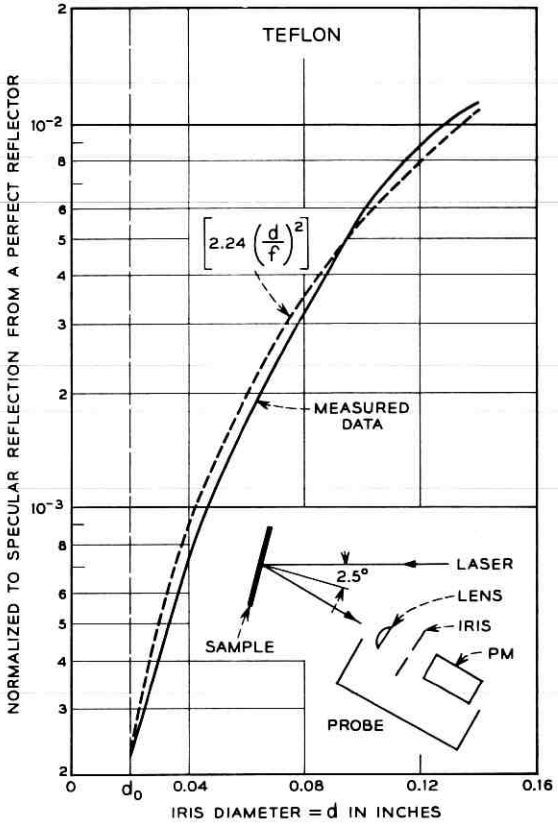


Fig. 7 — Scattering from Teflon in the direction of specular reflection.

With the above tests providing a degree of assurance in the probe's performance, scattering measurements in the specular direction as shown in Fig. 7, were made on the samples discussed in Figs. 6(a)–6(d). These data are shown as Figs. 8(a)–8(f). Here the solid curves represent the measured data and the dashed curves are calculations based upon the model that the measured power S_m* contains both a specularly reflected component R_{0m} and a scattered component $c(d/f)$,[2] where d is the diameter of the iris opening and f is the focal length of the lens. The constants of the equation $S_m = R_{0m} + c(d/f)^2$ are determined by fitting to the measured data. In reality, R_{0m} is the specular reflection coefficient for the surface. The parame-

* Normalized to the power specularly reflected from a perfect reflector.

ter, c, for the samples in Figs. 8(a)–8(f) is the relative scattered power per unit solid angle in the specular direction.

An examination of Figs. 8(a)–8(f) shows reasonable agreement between measured data and calculated values (always so for small values of d, where the curves were fitted); however, for those surfaces with a significant reflection coefficient (for example Fig. 8(a) for polished aluminum) the model predicts a larger value* than that measured for larger values of d. This effect is attributed to a preferred scattering in the specular direction by the individual surface facets. In Fig. 8(a), the point of divergence of the measured and fitted curves for polished aluminum occurs at $d = 0.08$ inches; the corresponding planar acceptance angle of the probe at this iris opening is $\beta = d/f = 0.04$ since f is two inches. Now it follows from elementary diffraction theory that if D is the dimension of an aperture (in this case a "reflecting" surface facet) and λ the wavelength of radiation, then the angle within which the radiation from this facet is concentrated is $\alpha = \lambda/D$. From the data in Fig. 8(a), one is therefore led to the conclusion (letting $\alpha = \beta$) that the facet dimension is $0.63\mu/0.04 = 16\mu$. Microscopic examination of the surface indicated an average facet size of about 40μ for the aluminum sample.

The specular reflection coefficients, R_{0m}, for all surfaces on which specular scattering measurements were made are tabulated in Table II.

The value given here for the specular reflection coefficient of a first surface aluminized mirror is about 10 per cent larger than the value usually quoted for aluminum film. Similarly there are differences in reflection coefficient for those samples listed with the (1.) and (2.) notation in Table II. Provided there are no polarization effects occurring at the surface facets, these specular reflection coefficients should be the same because the sample is merely rotated 90° about its axis in these two conditions.

Scattering in the specular direction having been determined, the scattering in the direction $\theta = 180°$ can now be plotted on the angular scattering curves, Figs. 6(a)–6(d); in those figures, the value at $\theta = 180°$ is joined to the angular scatter curve at $\theta = 175°$ by the dashes. This complete curve then represents the scattered energy as a function of θ. If circular symmetry obtains, a numerical integration yields the total scattered power.

* It should be noted that the calculated curves in some cases (e.g., Figs. 8(a), 8(d), and 8(e) exceed unity; this is because the square law dependence on iris opening assumes isotropically diffuse scattering from the facets which, of course, is not true if the facets have directivity.

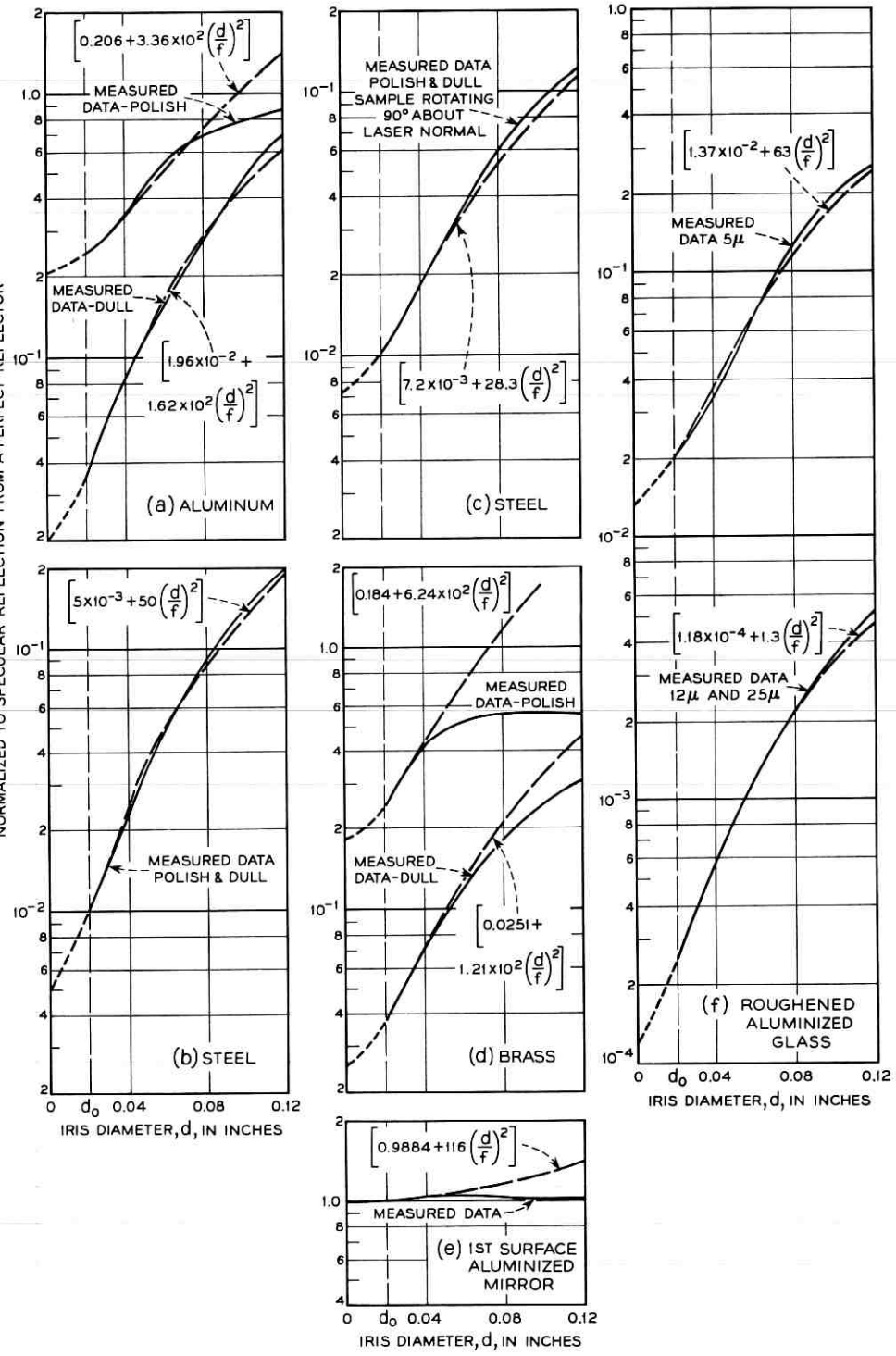


Fig. 8 — Scattering from various metallic surfaces in the direction of specular reflection.

TABLE II

Material	Specular Reflection Coefficient R_{0m}	Scattering in Specular Direction for an Iris Opening of d_0
Alum. Polish	0.21	0.034
Alum. Dull	0.02	0.016
Steel (polished & dull) (1)	5×10^{-3}	5×10^{-3}
Steel (polished & dull) (2)	7.2×10^{-3}	2.83×10^{-3}
Brass Polish	0.18	0.062
Brass Dull	0.025	0.012
1st Surface Mirror	0.988	0.012
5μ Roughened Glass Aluminized	0.014	6.3×10^{-3}
12μ & 25μ Roughened Glass Aluminized	1.2×10^{-4}	1.3×10^{-4}
Nickel Foil	0.97	0.032
Beryllium Copper-Polish (1)	0.013	0.031
Beryllium Copper-Dull (1)	5.5×10^{-3}	5.5×10^{-2}
Beryllium Copper-Polish (2)	0.016	0.014
Berylllium Copper-Dull (2)	6.5×10^{-3}	2.7×10^{-3}

Note: (1.) preferential scatter in plane of probe.
(2.) preferential scatter 90° to plane of probe.

VI. DISCUSSION OF THE SCATTERING MODEL FOR METALLIC SURFACES

First, consider unity power incident on an opaque (but otherwise arbitrary) surface. This power, reflected in the specular direction in proportion to the specular reflection coefficient, R_0, of the surface, is scattered over the hemisphere, or absorbed. Let $\sigma(\Omega)$ be the scattering coefficient of the surface in the direction Ω, then by conservation of energy,

$$1 = R_0 + \frac{1}{2\pi} \int_{2\pi} \sigma(\Omega)\, d\Omega + a \tag{7}$$

where a is the fraction of incident power absorbed by the surface and the integral term represents all the scattered power in the hemisphere.

Assuming azimuthal symmetry in the scattered power and letting $\sigma(\theta)$ be the scattering coefficient in the direction $\Omega(\theta, \sigma)$ then the power accepted by the probe looking toward the direction θ is (as in Section III) $S_m(\theta) = \sigma(\theta)\omega_0/4\pi$, the relative power measured by the probe looking toward the direction θ being designated by $S_m(\theta)$. Substituting for σ and $d\Omega$ in (7), one obtains

$$1 = R_0 + a + \frac{4\pi}{\omega_0} \int_{\pi}^{\pi/2} S_m(\theta) \sin\theta\, d\theta = R_0 + a + S_{TM} \tag{8}$$

where S_{TM} is the *total* scattered power. As previously stated, the acceptance angle of the focused probe is $\omega_0 = 10^{-3}$ steradians, thus when the

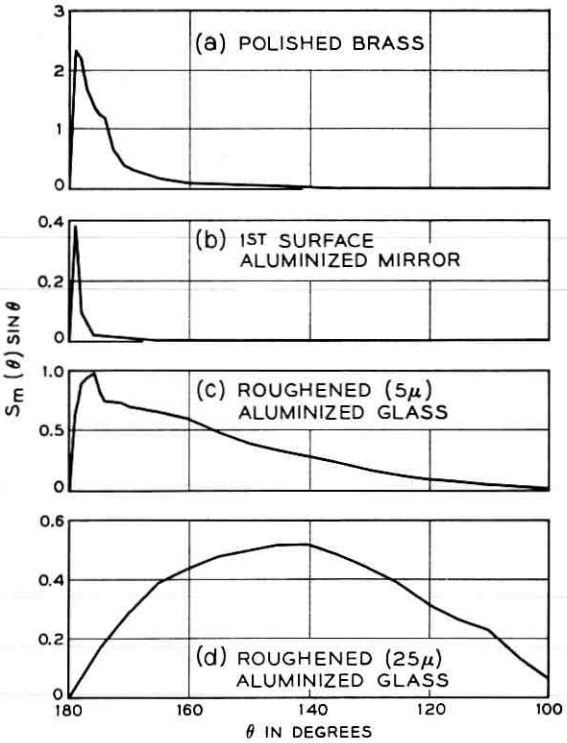


Fig. 9 — Integration curves for evaluation of total scattered power.

integral has been numerically evaluated, the fraction of power a absorbed by the surface can be determined.

The development of (8) assumed azimuthal symmetry whereas an examination of the scatter curves of Figs. 6(a)–6(c) shows these surfaces to be (in varying degrees) anisotropic; however, it is instructive to consider the curves representing the polished brass surface (Fig. 6(c))

TABLE III

Material	Measured Specular Reflection Coefficient R_{0m}	Measured Total Scattering Coeff. S_{TM}	Derived Absorption Coeff. (a)
Brass	0.18	0.44	0.38
1st Surface Mirror	0.988	0.01	0.0016
5μ Glass	1.4×10^{-2}	0.91	0.076
25μ Glass	1.2×10^{-4}	0.97	0.03

since the two curves* taken in the two azimuth planes are reasonably similar so that the symmetry is fairly good. A numerical integration of this curve yields a total scattered power of 0.44. The curve used for the numerical integration (the integrand of (8)) in this case is shown in Fig. 9(a).

In addition to the brass sample, numerical integrations of scattering have been made for the 1st surface mirror (Fig. 6(b)) and the roughened glass flats (Fig. 6(d)). Since the 12μ and 25μ roughened glass flats have similar scatter curves only the data for the 25μ glass sample was integrated. These data are given in Table III, along with the absorption coefficient a, as derived from (8).

The curves used for the numerical integration of the power scattered from these surfaces are shown in Fig. 9. Here one notices that the largest contribution from the mirror peaks around 179° and that the major contribution occurs in the first few degrees, whereas the curve for the 5μ glass has a decided peak occurring around 176° and has significant values out to fairly wide angles. The curve for 25μ glass peaks at about 140° since it is a very diffuse scatterer.

VII. CONCLUSIONS

Based upon the measurements, it would appear that an optical probe, with modifications to permit measuring scatter in the specular direction, is a useful tool for obtaining information on scattering from an arbitrary surface. The combination of angular scatter and specular scatter measurements have been successfully used in determining the total scattered power from an arbitrary surface. The method of measuring scatter in the specular direction has also produced values for the specular reflection coefficient and the absorption coefficient. Certain conclusions about facet size and the nature of the scattering process have been obtained from these data.

VIII. ACKNOWLEDGMENT

The interest and helpful suggestions of D. C. Hogg are greatly appreciated.

REFERENCES

1. Beckman, P., and Spizzichino, A., *The Scattering of Electromagnetic Waves from Rough Surfaces*, The Macmillan Co., 1963, p. 89.
2. Barrett, C. S., *Structure of Metals*, McGraw-Hill Book Co., 2nd Edition, Chap. 18, 1952.

* Curves such as these are examples of the $S_m(\theta)$ in (8).

Design Theory of Balanced Transistor Amplifiers

By K. KUROKAWA

(Manuscript received May 7, 1965)

This paper discusses the expected characteristics of balanced transistor amplifiers with symmetrical directional couplers.

Provided that pairs of transistors with similar characteristics can be selected from a given distribution, the input and output matches obtained with the balanced configuration are satisfactory over a ±10 per cent bandwidth with simple one-section lumped-constant LC directional couplers and over a ±40 per cent bandwidth (1.2 octaves) with one section distributed $\lambda/4$ couplers. For single-stage amplifiers, the decrease in gain is less than 0.1 db and the phase nonlinearities introduced by the couplers are about ±0.15° and ±0.6°, respectively, over the same bandwidths.

The requirements on the terminations which are connected to the couplers to absorb the transistor reflections are not stringent: VSWR's less than 1.4 should be acceptable. The noise measure of balanced amplifiers is calculated to be a weighted average of the noise measures of the two component amplifiers, plus a small term which vanishes when the couplers have 3-db coupling and the component amplifiers have identical gains. Gain compression takes place at a 3-db higher signal level compared with conventional single-ended designs, and the expected improvement in the third-order intermodulation is 9 db on the average.

In the final section, the cascade connection of identical balanced amplifiers is discussed. With typical microwave transistors, the input and output return losses for a multistage amplifier should be about 4.5 db worse than those for the individual single-stage amplifiers of which it is composed. The gain ripple introduced by the interactions between stages is also investigated in detail.

I. INTRODUCTION

In a previous paper,[1] the principles and experimental results of an L-band balanced transistor amplifier have been discussed in which each

stage consists of two electrically similar transistors whose inputs and outputs are combined through 3-db directional couplers. Due to wide distributions in the characteristics of present microwave transistors, simultaneous realization of flat gain and good impedance matching is difficult to obtain with conventional single-ended designs unless, for instance, isolators are employed. On the other hand, as long as pairs of transistors with similar characteristics can be chosen from a given distribution, the balanced design offers good input and output impedance matches as well as smooth gain and phase characteristics, all simultaneously. Since the impedance matches are important in microwave systems, the balanced design will be useful for some time, until the distribution of transistor characteristics becomes so tight that a conventional single-ended design can easily provide good matches and smooth gain simultaneously.

While the theory given in the paper mentioned above should be adequate for general purposes, it may not be satisfactory for the actual design of balanced transistor amplifiers. The theory neglected interactions between the reflections which were introduced to explain the mismatches at the input and output ports of the transistors. It also assumed ideal 3-db coupling of the directional couplers for the entire frequency band of interest. The former shortcoming can be avoided by employing scattering matrices in the discussion. This paper is intended to supplement the previous one by presenting an improved theory which enables us to discuss the effect of the coupling variation on amplifier characteristics without resorting to too complicated mathematics. The noise performance and intermodulation characteristics are also included. In the final section, interactions between stages — when connected in cascade — are discussed in detail. Although wider bandwidths may be obtained by using couplers having characteristics slightly different from one another (e.g. stagger-tuning), for simplicity this paper assumes the use of identical couplers for both single and multistage amplifiers.

II. REVIEW OF DIRECTIONAL COUPLER PRINCIPLES

A directional coupler is a matched four-port network with zero coupling between conjugate ports. Let us consider a symmetrical directional coupler with two planes of symmetry as shown in Fig. 1. Because of the symmetry and by definition, the scattering matrix must have the form

$$S = \begin{bmatrix} 0 & \alpha & \beta & 0 \\ \alpha & 0 & 0 & \beta \\ \beta & 0 & 0 & \alpha \\ 0 & \beta & \alpha & 0 \end{bmatrix} \tag{1}$$

If a network is lossless, the scattering matrix has to satisfy

$$S^{+}S = I \tag{2}$$

where $+$ indicates the transposed conjugate matrix and I the unit matrix. From this, two constraints between α and β of a lossless symmetrical coupler are obtained:

$$|\alpha|^2 + |\beta|^2 = 1, \qquad \alpha^*\beta + \beta^*\alpha = 0. \tag{3}$$

The first equation specifies a relation between magnitudes and the second one indicates that α and β must be 90° out of phase. Thus, α and β must be expressible in terms of two real quantities t and φ,

$$\alpha = \sqrt{1 - t^2}\, j \exp(-j\varphi), \qquad \beta = t \exp(-j\varphi). \tag{4}$$

That is, if a unit wave is incident on port 1, the output waves from ports 2 and 3 are given by $\sqrt{1 - t^2}\, j \exp(-j\varphi)$ and $t \exp(-j\varphi)$, respectively, and port 4 has no output.

There is a class of symmetrical junctions which acts as directional couplers independent of the frequency. Let us consider two of them as examples: a lumped constant directional coupler, and a distributed

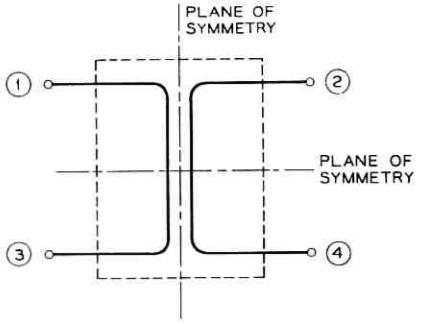


Fig. 1 — Schematic diagram of symmetrical directional coupler.

transmission line directional coupler $\lambda/4$ long at center frequency of operation.

For the lumped-constant directional couplers (Appendix) with a common inductance L from ports 1 and 2 to ports 3 and 4 and with a capacitance C between ports 1 and 2 or 3 and 4,

$$t = \frac{1}{\sqrt{1 + \zeta^2}}, \qquad \varphi = \tan^{-1} \zeta \tag{5}$$

where $\zeta = \omega L/Z_o$ and the characteristic impedance $Z_o = \sqrt{L/C}$. When $\zeta = 1$, $t^2 = 1 - t^2 = 0.5$ and 3-db coupling is obtained. Figs. 2 and 3 show t and φ vs the normalized frequency f/f_o where f_o is the frequency for ζ being 1.

For the distributed directional coupler[2]

$$t = \sqrt{\frac{1 - k^2}{1 - k^2 \cos^2 \theta}}, \qquad \varphi = \tan^{-1} \frac{1}{\sqrt{1 - k^2}} \tan \theta \tag{6}$$

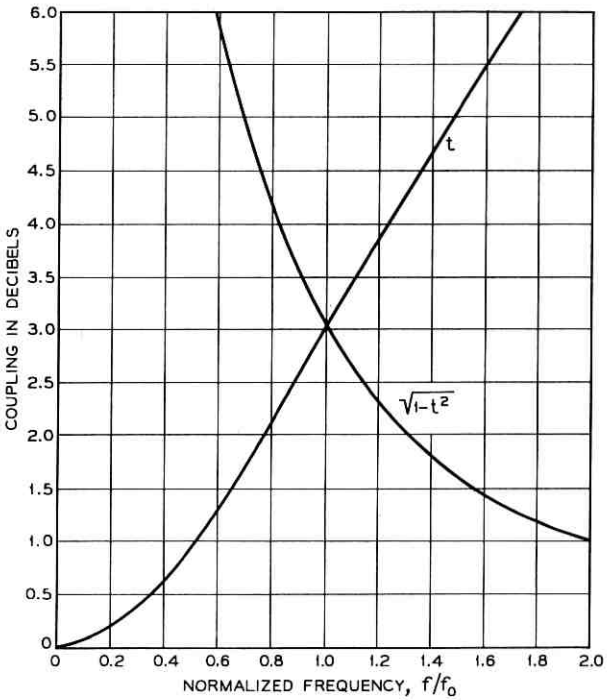


Fig. 2 — Coupling vs normalized frequency of lumped-constant LC directional coupler.

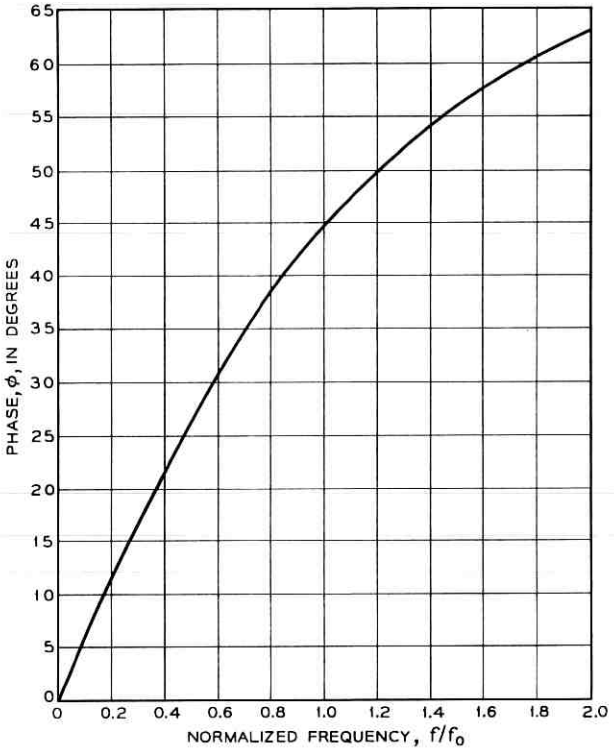


Fig. 3 — Phase vs normalized frequency of lumped-constant LC directional coupler.

where k is the coupling factor in the theory of coupled transmission lines and θ is the electrical length of the coupled region. In terms of the even- and odd-mode characteristic impedances Z_{oe} and Z_{oo}, the characteristic impedance Z_o and the coupling factor k are given by

$$Z_o = \sqrt{Z_{oe}Z_{oo}}, \qquad k = \frac{Z_{oe} - Z_{oo}}{Z_{oe} + Z_{oo}}, \tag{7}$$

respectively. Figs. 4 and 5 give t and φ vs the normalized frequency f/f_o and f_o is the frequency for which $\theta = \pi/2$ or 90°.

III. SCATTERING MATRIX OF ONE-STAGE BALANCED AMPLIFIER

Let us consider the configuration shown in Fig. 6 where two transistors, a and b, are connected by two directional couplers in which ports 3 and 4 are crossed over (as compared to Fig. 1). Due to the lack

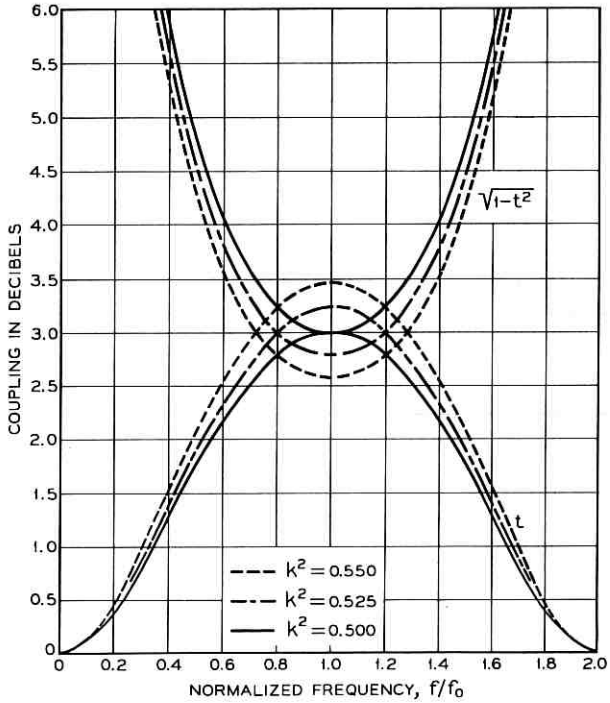


Fig. 4 — Coupling vs normalized frequency of one-section distributed couplers. k: coupling factor.

of coupling between conjugate arms (1–4 and 2–3 in Fig. 1) of the couplers, the components of the over-all scattering matrix between ports 1 and 2, are easily calculated. They are:

$$\begin{aligned} S_{11} &= e^{-2j\varphi}[t^2 S_{11}(a) - (1 - t^2)S_{11}(b)] \\ S_{21} &= je^{-j2\varphi}t\sqrt{1 - t^2}\,[S_{21}(a) + S_{21}(b)] \\ S_{12} &= je^{-j2\varphi}t\sqrt{1 - t^2}\,[S_{12}(a) + S_{12}(b)] \\ S_{22} &= e^{-2j\varphi}[t^2 S_{22}(b) - (1 - t^2)S_{22}(a)]. \end{aligned} \tag{8}$$

The subscripts 1 and 2 refer to the input and output ports, respectively, and $S_{11}(a)$, $S_{11}(b)$ etc., are the scattering matrix components of transistors a and b including their surrounding circuits, i.e., the scattering matrix components of the component amplifiers a and b, respectively. When the coupling is 3 db ($t^2 = 0.5$) and the two component amplifiers are similar in their characteristics,

$$S_{ij}(a) \approx S_{ij}(b)$$

and hence

$$|S_{11}| \approx 0, \qquad |S_{22}| \approx 0, \qquad S_{21} \approx je^{-j2\varphi}S_{21}(a) \approx je^{-2j\varphi}S_{21}(b).$$

This means that the input and output ports of the balanced amplifier are well matched and the gain is approximately equal to that of either

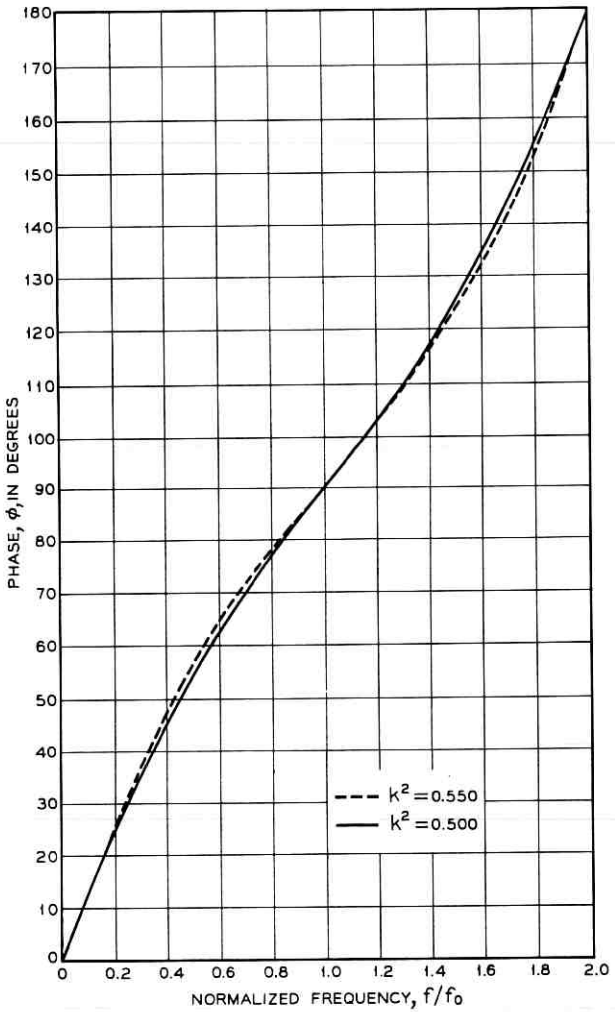


Fig. 5 — Phase vs normalized frequency of one-section distributed couplers.

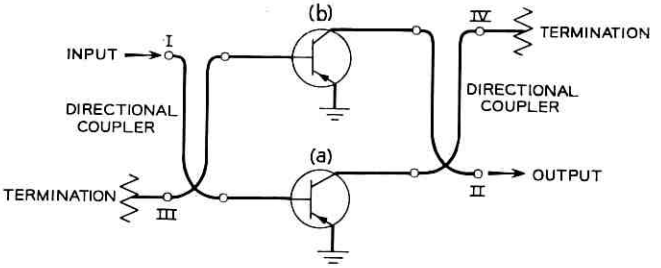


Fig. 6 — Schematic diagram of one-stage balanced transistor amplifier.

component amplifier. Because of the term, exp $(-j2\varphi)$, the phase of the balanced amplifier is affected by the 3-db couplers. This will be discussed in detail in Section V. The reflections from the transistors are absorbed in the terminations connected to the couplers as we shall discuss in Section VI.

When the transistors are dissimilar, but with the coupling still 3 db,

$$| S_{11} | = \tfrac{1}{2} | S_{11}(a) - S_{11}(b) |, \quad | S_{22} | = \tfrac{1}{2} | S_{22}(a) - S_{22}(b) |$$

and

$$| S_{21} | = \tfrac{1}{2} | S_{21}(a) + S_{21}(b) |.$$

That is, the input and output reflections are reduced to half of the vector differences of the corresponding reflections of the transistors, and the gain is given by the vector average of the two gains.

In this section, the coupling of the directional couplers has been assumed to be 3 db. In Section IV, we shall investigate the effect on the amplifier characteristics when the coupling is not 3 db.

IV. COUPLING OF THE DIRECTIONAL COUPLERS AND AMPLIFIER CHARACTERISTICS

First, let us assume that the two component amplifiers are similar in their characteristics. Then, from (8), we have

$$\begin{aligned} | S_{11} | &\approx | 2t^2 - 1 | \, | S_{11}(a) | \\ | S_{22} | &\approx | 2t^2 - 1 | \, | S_{22}(a) | \\ | S_{21} | &\approx | 2t\sqrt{1 - t^2} | \, | S_{21}(a) |. \end{aligned} \tag{9}$$

If $| 2t^2 - 1 |$ is given in terms of loss in db and $| S_{11}(a) |$ or $| S_{22}(a) |$

in terms of return loss in db, then the addition of these figures gives the corresponding return loss of the balanced amplifier in db. Similarly, if the gain of the component amplifiers, $|S_{21}(a)|$, is expressed in db and $|2t\sqrt{1-t^2}|$ in terms of loss in db, then the difference of these two figures gives the gain of the balanced amplifier in db. Fig. 7 shows the losses $|2t^2-1|$ and $|2t\sqrt{1-t^2}|$ in db vs the coupling loss t in db. From Fig. 7 we see, for instance, that if the input and output VSWR's of the component amplifiers are better than 2 (return loss 10 db), the coupling loss t can deviate as much as -0.4 db and $+0.5$ db from 3 db before the VSWR's of the balanced amplifier become worse than 1.07 (return loss 30 db) and that the decrease of the gain due to the directional couplers from that of the component amplifier is less than 0.1 db. The above deviation allows ± 10 per cent and ± 40 per cent frequency bandwidths for the lumped-constant and the distributed ($k^2 = 0.550$) couplers, respectively.

Next, let us consider the case where the two component amplifiers have different characteristics. In this case, from (8), we have

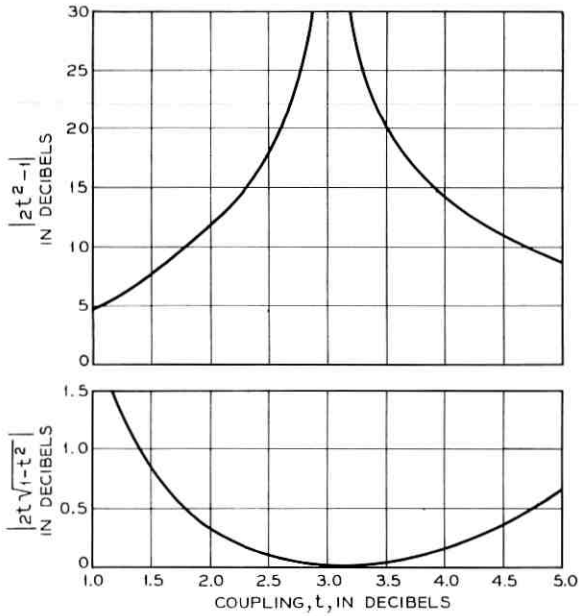


Fig. 7 — Improvement in return loss $|2t^2-1|$ in db and decrease in gain $|2t\sqrt{1-t^2}|$ in db due to the directional couplers in balanced design.

$$
\begin{aligned}
|S_{11}| &= \left| (2t^2 - 1)\frac{S_{11}(a) + S_{11}(b)}{2} + \frac{S_{11}(a) - S_{11}(b)}{2} \right| \\
|S_{22}| &= \left| (2t^2 - 1)\frac{S_{22}(a) + S_{22}(b)}{2} - \frac{S_{22}(a) - S_{22}(b)}{2} \right| \\
|S_{21}| &= |2t\sqrt{1 - t^2}| \left| \frac{S_{21}(a) + S_{21}(b)}{2} \right|.
\end{aligned} \tag{10}
$$

The gain expression in (10) is the same as that in (9) except $S_{21}(a)$ is replaced by the mean vector between $S_{21}(a)$ and $S_{21}(b)$. Therefore, using the magnitude of the mean vector and Fig. 7, the expected gain of the balanced amplifier can be easily calculated. The reflections $|S_{11}|$ and $|S_{22}|$ have two terms each: one becomes small when the coupling approaches 3 db and the other is independent of the coupling. The magnitude of the first term can be evaluated by using Fig. 7 as we have done before. Now, however, the mean reflection from the two amplifiers is used instead of the same reflection from the component amplifiers. The magnitude of the second term is half of the difference between the reflections from the two component amplifiers. In order to get the resultant reflection, however, a vectorial addition of these two terms is necessary. For more clear understanding of the situation, the following is an additional way of viewing the same problem.

Rearranging the first two equations in (10), we have

$$
\begin{aligned}
|S_{11}| &= |(2t^2 - 1)S_{11}(a) - (1 - t^2)\Delta_1| \\
|S_{22}| &= |(2t^2 - 1)S_{22}(a) + t^2\Delta_2|,
\end{aligned} \tag{11}
$$

where Δ_1 and Δ_2 are given by

$$
\Delta_1 = S_{11}(b) - S_{11}(a), \qquad \Delta_2 = S_{22}(b) - S_{22}(a).
$$

Suppose that the coupling t and $S_{11}(a)$ are specified and that the resultant reflection $|S_{11}|$ is required to be smaller than a certain magnitude, $|S_{11}|_{\max}$. Then, referring to Fig. 8, the tip of $-(1 - t^2)\Delta_1$, drawn from $(2t^2 - 1)S_{11}(a)$, must lie inside a circle centered at the origin and of radius $|S_{11}|_{\max}$. Expanding the figure by a factor of $-1/(1 - t^2)$, the tip of Δ_1, drawn from $-(2t^2 - 1)S_{11}(a)/(1 - t^2)$, is seen to be inside a circle centered at the origin and of radius $|S_{11}|_{\max}/(1 - t^2)$. Now translate the whole figure until the tail of Δ_1 falls on the point $S_{11}(a)$. Then, $S_{11}(b)$ is seen to be inside a circle centered at $t^2S_{11}(a)/(1 - t^2)$ and of radius $|S_{11}|_{\max}/(1 - t^2)$ on the Smith chart. Similarly, if the maximum allowable value of $|S_{22}|$ is given by $|S_{22}|_{\max}$, $S_{22}(b)$ must be inside a circle centered at $(1 - t^2)S_{22}(a)/t^2$ and of radius $|S_{22}|_{\max}/t^2$.

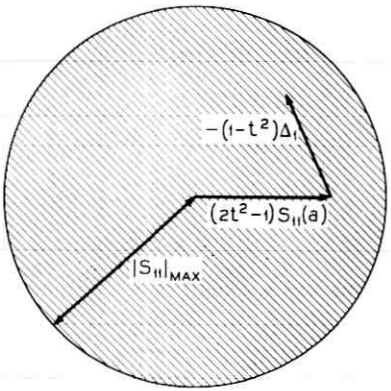


Fig. 8 — Vector diagram for the relation $| S_{11} | \leqq | S_{11} |_{max}$.

As t^2 increases, the area in which $S_{11}(b)$ should be located also increases; however, the area for $S_{22}(b)$ decreases. The best compromise is obtained when the coupling is 3 db. Here, one might ask the following question: If there is no requirement for $| S_{22} |$, should one make t^2 as large as possible in order to achieve the required matching more easily? The answer is, generally, no. Since $S_{11}(a)$ is usually larger than $| S_{11} |_{max}$, as t^2 approaches 1, the center $t^2 S_{11}(a)/(1 - t^2)$ moves faster than the increase in the radius $| S_{11} |_{max}/(1 - t^2)$. The circle therefore ceases to cover the area where the transistor distribution for S_{11} is dense and it becomes harder to find a proper transistor which gives the required $S_{11}(b)$. In this argument, if $S_{11}(a)$ is always smaller than $| S_{11} |_{max}$, t^2 could obviously be 1. However, if the reflection from the component amplifier a is already smaller than the required value there is no reason for using the balanced configuration and a second amplifier. A similar argument holds for the output match as well.

V. PHASE LINEARITY

It is obvious from (8) that the phase linearity of the balanced amplifier depends on the phase characteristics of the directional couplers as well as on the phase linearity of the component amplifiers. The φ's of the couplers were discussed in Section II and are shown in Figs. 3 and 5. From these figures, the phase nonlinearity introduced by the directional couplers can be estimated. However, in precise applications the over-all phase linearity required is often within a few degrees and if several stages in cascade are employed to obtain a desired gain, the phase linearity required for each stage would be within a fraction of one degree. In such a case, the Taylor expansion of 2φ around $f = f_o$ should give a

better estimate of the nonlinearity introduced by the couplers. For the lumped constant directional couplers,

$$\begin{aligned} 2\varphi &= \tfrac{1}{2}\pi + \chi - \tfrac{1}{2}\chi^2 + \tfrac{1}{6}\chi^3 + \cdots \qquad (2\varphi, \text{ in radians}) \\ &= 90 + 57.3\chi - 28.7\chi^2 + 9.6\chi^3 + \cdots \qquad (2\varphi, \text{ in degrees}) \end{aligned} \tag{12}$$

where

$$\chi = (f - f_o)/f_o. \tag{13}$$

For the distributed couplers $\lambda/4$ long at f_o,

$$\begin{aligned} 2\varphi &= \pi + \pi\sqrt{1 - k^2}\,\chi + \frac{k^2\sqrt{1 - k^2}}{12}\pi^3\chi^3 \\ &\quad + \frac{k^2(3k^2 - 1)\sqrt{1 - k^2}}{240}\pi^5\chi^5 + \cdots \quad (2\varphi, \text{ in radians}) \\ &= 180 + 180\sqrt{1 - k^2}\,\chi + 148\,k^2\sqrt{1 - k^2}\,\chi^3 \\ &\quad + 73k^2(3k^2 - 1)\sqrt{1 - k^2}\,\chi^5 + \cdots \quad (2\varphi, \text{ in degrees}). \end{aligned} \tag{14}$$

For instance, the deviations of the lumped-constant couplers from phase linearity at $\chi = f/f_o - 1 = 0.1$, 0.2 and 0.3 are 0.3°, 1.2° and 2.6° respectively, and of the order of 0.04°, 0.3° and 1° for the distributed couplers. These deviations are measured from the straight line tangential to the 2φ curve at $f = f_o$. If the reference line is redrawn so that the maximum deviation becomes the smallest, these figures will be about one-half of those mentioned above for the lumped-constant couplers and about one-fourth for the distributed couplers. Thus, over the bandwidths for which the input and output VSWR's are less than 1.07 as discussed before, the phase nonlinearities introduced are of the order of $\pm 0.15°$ and $\pm 0.6°$, respectively.

VI. EFFECT OF IMPERFECT TERMINATIONS

In order to investigate in detail the role of the terminations connected to the couplers, let us first consider the balanced amplifier as a four-port network rather than a two-port network as we have done so far. The ports are numbered by Roman numerals as shown in Fig. 6. Since there is no coupling between conjugate ports of the couplers, the scattering matrix of the four-port network can again be easily calculated. S_{11}, S_{21}, S_{12}, and S_{22} are the same as given in (8). S_{43} and S_{34} are equal to S_{21} and S_{12} respectively. S_{33} and S_{44} are the same as S_{11} and S_{22}, respectively, except that the component amplifier designations a and b are interchanged. The others are given by

$$
\begin{aligned}
S_{13} &= S_{31} = je^{-2\varphi}t\sqrt{1-t^2}\,[S_{11}(a) + S_{11}(b)] \\
S_{24} &= S_{42} = je^{-2j\varphi}t\sqrt{1-t^2}\,[S_{22}(a) + S_{22}(b)] \\
S_{23} &= e^{-2j\varphi}[t^2S_{21}(b) - (1-t^2)S_{21}(a)] \\
S_{32} &= e^{-2j\varphi}[t^2S_{12}(b) - (1-t^2)S_{12}(a)] \\
S_{14} &= e^{-2j\varphi}[t^2S_{12}(a) - (1-t^2)S_{12}(b)] \\
S_{41} &= e^{-2j\varphi}[t^2S_{21}(a) - (1-t^2)S_{21}(b)].
\end{aligned} \tag{15}
$$

When a wave of unity power is incident to port I, it is split by the input coupler into two, t^2 and $(1-t^2)$, arriving at the component amplifiers a and b, respectively. The reflections there are given by $t^2 \mid S_{11}(a) \mid^2$ and $(1-t^2) \mid S_{11}(b) \mid^2$. Of these, only $\mid S_{11} \mid^2$ comes back to port I and the rest of the power

$$
\begin{aligned}
t^2 \mid S_{11}(a) \mid^2 + (1-t^2) \mid S_{11}(b) \mid^2 - \mid S_{11} \mid^2 \\
= t^2(1-t^2) \mid S_{11}(a) + S_{11}(b) \mid^2
\end{aligned} \tag{16}
$$

(which is exactly equal to $\mid S_{31} \mid^2$), goes to port III. Thus, most of the reflected power from the component amplifiers goes to port III and is absorbed there. A similar argument holds for the output port. These are the reasons why the balanced configuration gives good matches at both ends.

Next, let us investigate how critical the matches are for these terminations. Indicating the reflection coefficients of the terminations by r_3 and r_4 (the subscripts refer to port numbers) and drawing the signal flow graph as shown in Fig. 9, the reflection to port I, S_{11}' can be written down by inspection.

$$
S_{11}' = S_{11} + \frac{\begin{array}{l} r_3S_{13}S_{31}(1 - r_4S_{44}) + r_4S_{14}S_{41}(1 \\ \quad - r_3S_{33}) + r_3r_4(S_{13}S_{34}S_{41} + S_{14}S_{43}S_{31}) \end{array}}{1 - r_3S_{33} - r_4S_{44} - r_3r_4S_{43}S_{34} + r_3r_4S_{33}S_{44}}. \tag{17}
$$

Neglecting higher order terms, S_{11}' can be approximated by

$$
S_{11}' \approx S_{11} + r_3S_{13}S_{31} + r_4S_{14}S_{41}. \tag{18}
$$

The first term on the right-hand side represents the reflection when $r_3 = r_4 = 0$, the second term the reflection due to r_3 and the last term due to r_4. The neglected terms represent the contribution due to multireflections between the terminations and they are in general so small compared with the terms given above that their omission is readily justified.

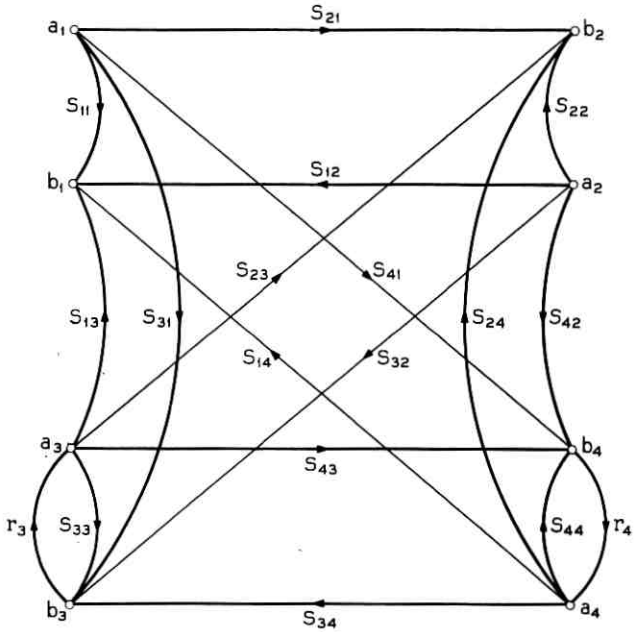


Fig. 9 — Signal flow graph for one-stage balanced amplifier with imperfect terminations.

Similarly, S_{22}' and S_{21}' are approximately given by

$$\begin{aligned} S_{22}' &\approx S_{22} + r_4 S_{24} S_{42} + r_3 S_{23} S_{32} \\ S_{21}' &\approx S_{21} + r_3 S_{23} S_{31} + r_4 S_{24} S_{41}. \end{aligned} \tag{19}$$

Typical values for the magnitude of the coefficients of r_3 and r_4 appearing in S_{11}' and S_{22}' are of the order of -20 db and those appearing in S_{21}' are of the order of -25 db or less. Therefore, if $|r_3|$ and $|r_4|$ are both less than -15 db (SWR $\leqq$ 1.4), then the effect of imperfect terminations on the amplifier characteristics must be negligibly small. In conclusion, the required matches for the terminations are in general not so stringent; SWR's of less than 1.4 are usually acceptable.

VII. NOISE PERFORMANCE

Noise performance of an amplifier is evaluated by the actual noise measure.[3] It is defined by

$$M = \frac{F - 1}{1 - (1/G)} \tag{20}$$

where G is the transducer gain and F is the noise figure of the amplifier, including the contribution of the noise power originating in and reflected back to the output load. When the input and output of the amplifier are matched, a number of amplifiers with identical characteristics can be connected in cascade, making the total gain very high. The excess noise figure of the high gain amplifier is then given by M itself. However, when the input and output are not matched, we have only to insert isolators between the stages in order to reach the same interpretation for M. For each amplifier there is an optimum value, M_{opt}, of M which can be achieved by a lossless imbedding but cannot be surpassed by any passive transformation of the amplifier. The noise measure itself is a dimensionless number.

Now, let us consider the balanced amplifier. The terminations III and IV connected to ports III and IV, respectively, are assumed to be matched. Furthermore, the noise temperatures of the terminations as well as of the load are assumed to be 290°K. For the ideal case where the coupling of the directional couplers is 3 db and the characteristics of the two component amplifiers are identical, the noise originating in each component amplifier is split into two. Only a half of the total power goes to the output load, with the other half going to termination IV. Since there are two component amplifiers, the output load receives the same noise power as in the single-ended design. The noise originating in termination III is amplified but absorbed in termination IV and none of it comes out to the load. The noise originating in the load and reflected back from the component amplifiers goes to termination IV and does not come back to the load. However, noise power originating in termination IV goes into the load. This power is exactly equal to the noise power originating in and reflected back to the load ($T = 290°$) in the single-ended design. As a result, the noise measure M of the one-stage balanced amplifier in this ideal case is equal to the noise measure of either component amplifier.

When a transistor is unconditionally stable, by inserting a proper lossless circuit at the input and a matching circuit at the output, the optimum value M_{opt} for the transistor can be achieved. Therefore, the component amplifiers can have M_{opt} in this way, which means that the balanced amplifier can also give M_{opt}. It is worth noting that this realization of M_{opt} does not deteriorate the input matching of the balanced amplifier. In general, this is not the case for single-ended designs.

Next, let us consider the case where the coupling is not necessarily 3 db and the component amplifiers have different characteristics. The assumptions for the terminations and the load remain the same as be-

fore. The excess noise output to the load is given by

$$\begin{aligned} N = {} & \{(F_a - 1)kTB \mid S_{21}(a) \mid^2 - kTB \mid S_{22}(a) \mid^2\}(1 - t^2) \\ & + \{(F_b - 1)kTB \mid S_{21}(b) \mid^2 - kTB \mid S_{22}(b) \mid^2\}t^2 \\ & + kTB \mid S_{23} \mid^2 + kTB \mid S_{24} \mid^2 + kTB \mid S_{22} \mid^2, \end{aligned} \tag{21}$$

where F_a and F_b are the noise figures of the component amplifiers a and b, respectively. The first term on the right-hand side of (21) represents the output noise originating in component amplifier a, the second one in component amplifier b, the third one in termination III, the fourth one in termination IV, and the last one represents the noise originating in and reflected back to the load. Combination of (8), (15) and (21) together with (20) gives the noise measure M of the balanced amplifier as follows:

$$M = \frac{M_a(\mid S_{21}(a) \mid^2 - 1)(1 - t^2) + M_b(\mid S_{21}(b) \mid^2 - 1)t^2 + \mid S_{23} \mid^2}{\mid S_{21} \mid^2 - 1}. \tag{22}$$

Thus, M is a weighted average of the noise measures M_a and M_b of the component amplifiers plus a small term which comes in because of termination III. To make this additional term small, from (15) $t^2 S_{21}(b)$ and $(1 - t^2)S_{21}(a)$ should be close to each other. When the coupling is 3 db, this means that the two component amplifiers should have approximately equal gains. Thus, we see that for the balanced design, a pair of transistors should be selected on the basis of close S_{11}, S_{22} and S_{21}; the first two being necessary for good matches and the last for low noise (although, in practice, the last requirement is not stringent at all).

VIII. GAIN COMPRESSION AND INTERMODULATION

Since each transistor handles only one-half of the signal power, it begins to saturate at a 3-db higher signal level thus improving the gain compression and intermodulation characteristics. The type of intermodulation of most concern in broadband amplifiers with multiple frequency channels is usually one in which two strong signals of frequency f_A and f_B produce third order intermodulation signals at frequencies $2f_A - f_B$ and $2f_B - f_A$, also within the passband of the amplifier, where they might interfere with wanted weak signals. Since the signal level to each transistor is 3-db lower, the third order intermodulation output from each transistor must be 9-db lower. Thus, if the two intermodulation outputs are in phase, a resultant output of 6-db below that of the single-ended design is expected for the balanced amplifier.

However, the magnitude of the third order intermodulation output varies between transistors at microwave frequencies, even if the transistor characteristics for the signal frequency are quite similar. This suggests that the phase of the intermodulation output might also be random. In this case, the resultant output of the balanced amplifier is expected to be 9-db lower on the average, instead of 6 db. This conclusion is strongly supported by experimental results on 18 different pairs of transistors.[4]

IX. CASCADE CONNECTION

So far we have discussed only single-stage balanced amplifiers. In this section, let us consider the interactions between stages when connected in cascade. Since the outgoing wave of the nth stage output port is equal to the incoming wave of the $(n + 1)$th stage input port, the signal flow graph of a multistage amplifier becomes something like Fig. 10, where $S_{11}(n)$ etc., are the scattering matrix components of the

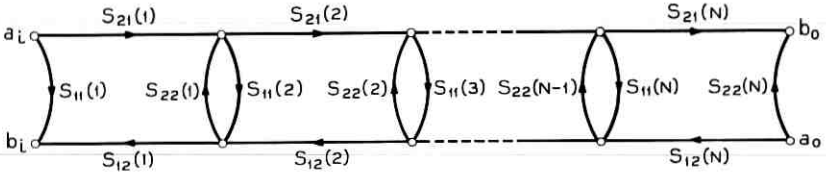


Fig. 10 — Signal flow graph of multistage amplifier.

nth stage and a_i, b_i, a_o and b_o are the incident and outgoing waves at the input (subscript i) and the output (subscript o) ports of the cascaded amplifier, respectively. Inspecting this signal flow graph, the components of the over-all scattering matrix can be obtained. First, taking a three-stage amplifier as an example, let us consider the input reflection S_{11} of the amplifier. It is given by

$$S_{11} = S_{11}(1) + \frac{S_{11}(2)S_{21}(1)S_{12}(1)}{\Delta}\{1 - S_{22}(2)S_{11}(3)\} + \frac{S_{11}(3)S_{21}(1)S_{12}(1)S_{21}(2)S_{12}(2)}{\Delta}, \quad (23)$$

where

$$\Delta = 1 - S_{22}(1)S_{11}(2) - S_{22}(2)S_{11}(3) - S_{22}(1)S_{21}(2)S_{11}(3)S_{12}(2) + S_{22}(1)S_{11}(2)S_{22}(2)S_{11}(3). \quad (24)$$

Since $| S_{22}(n)S_{11}(n+1) |$ is small compared with unity for most practical balanced amplifiers and since $\rho = | S_{21}(n)S_{12}(n) |$ is of the order of 0.4 for typical transistors, S_{11} can be approximated by

$$S_{11} \approx S_{11}(1) + S_{11}(2)S_{21}(1)S_{12}(1) + S_{11}(3)S_{21}(1)S_{12}(1)S_{21}(2)S_{12}(2). \tag{25}$$

The first, second, and third terms on the right-hand side of (25) represent the contributions to S_{11} by the reflections from the first, second and third stages respectively. The effect of later stages is seen to be reduced by the buffer action of the previous stages indicated by $S_{21}(n)S_{12}(n)$. For instance, when $\rho = | S_{21}(n)S_{12}(n) |$ is approximately equal to 0.4, the contribution of the third stage to the over-all mismatch is reduced to only $\rho^2 = 0.16$ times the original reflection. When the frequency is changed, the phase of $S_{21}(n)S_{12}(n)$ as well as that of $S_{11}(n)$ changes. Therefore, the vectors representing the successive terms on the right-hand side of (25) are expected to rotate with successively increasing speeds. At some frequencies, they tend to cancel each other and at other frequencies they tend to add up. Thus, if the reflection of each stage is of the same order of magnitude and $\rho = | S_{21}(n)S_{12}(n) | \approx 0.4$, then for the three-stage amplifier, $1 + \rho + \rho^2 \approx 1.56$ times as large reflection as the single stage should be expected (or 4 db worse return loss). Similarly, for a multistage amplifier with a large number of stages, $1 + \rho + \rho^2 + \cdots = 1/1 - \rho \approx 1.67$ times as large reflection should be anticipated (or 4.5 db worse return loss). The output reflection S_{22} can be discussed in a similar manner.

Next, let us consider the gain S_{21} . S_{21} of the three-stage amplifier is is given by

$$S_{21} = \frac{S_{21}(1)S_{21}(2)S_{21}(3)}{\Delta}, \tag{26}$$

where Δ is given by (24).

Since factors other than the effect of Δ being different than one were dominant, Δ could be approximated by unity for the discussion of S_{11} . However, for discussing S_{21} , Δ has to be investigated in detail. When each stage is well matched, Δ is unity and the over-all gain is the product of the gain of each stage as expected. The effect on the gain of the interaction between stages comes from the various terms in Δ. The second and third terms on the right-hand side of (24) represent the effect of the interaction between the adjacent stages. The fourth term shows the effect of the interaction between the first and the third stages through

some buffer action of the second stage. The last term gives the higher order interaction which for well designed balanced amplifiers can usually be neglected. Since the phases as well as the magnitudes of the $S_{ij}(n)$'s change with frequency, the interactions between stages introduce ripples in the gain vs frequency curve. When $|S_{11}(n)|$ and $|S_{22}(n)|$ are of the order of 0.1 (or 20-db return loss) the magnitudes of the ripples introduced by the second and third terms in Δ are of the order of ± 0.1 db. Therefore, the expected value of the resultant is ± 0.14 db (or ± 0.2 db for the worst case). The magnitude of the ripple due to the fourth term is smaller by the factor of $|S_{21}(2)S_{12}(2)|$. However, the phase of $S_{21}(2)S_{12}(2)$ increases the speed with which the vector rotates with frequency. This rapid variation is sometimes troublesome. To reduce the repetition rate of the gain variation with frequency, the equivalent electrical length of each stage should be made as small as possible. Also, transistors with high reverse loss help reduce the magnitude of the rapid ripple.

For an N-stage amplifier, the corresponding Δ includes $N - 1$ terms representing the interactions between the adjacent stages, $N - 2$ terms representing those between the nth and $n + 2$nd stages and so forth — with additional terms representing various higher order interactions. The contribution from each group to the gain ripple is proportional to $\sqrt{N-1}$, $\sqrt{N-2}\,\rho$ and so forth, provided that all stages have similar characteristics. Although the speed of the rotation with frequency increases successively, the magnitude of the vector representing each group diminishes rapidly and practically no interaction between the stages beyond 3 stages away from each other can be observed when $\rho \approx 0.4$.

Since each balanced stage is, in practice, well matched at both ends, the noise performance and intermodulation characteristics of a multistage amplifier with identical stages are clear from the discussions of the single-stage amplifier. However, because the main noise contribution comes from the first few stages and the contribution to the compression or intermodulation comes from the last few stages, it may be advisable not to use an identical design for all stages. Instead, the first few stages can be designed for best noise performance and the last few stages for best compression characteristics. This can usually be done by changing only the transistor dc bias circuit.

In large scale production, when identical circuits are to be used for the first several stages of each amplifier, the following procedure of selecting transistor pairs gives the best noise performance on the average. First, obtain the actual noise measure M of all transistors in a standard component amplifier and classify them into several groups of increasing

M, each group containing twice as many transistors as the number of amplifiers to be built. Then select pairs of transistors from each group separately on the basis of similar scattering matrices and use first, each pair from the best group (lowest M) in the first stage of each amplifier, next use those from the second group in the second stage and so forth. The pairs from the last group with poor M's must be used in the later stages, whose noise contributions are insignificant. A similar procedure can be applied to the selection of transistor pairs for best compression characteristics. Here, of course, each pair from the best compression group is used in the last stage of each amplifier.

X. ACKNOWLEDGMENTS

Acknowledgments are due to R. S. Engelbrecht who originally suggested the balanced transistor amplifiers and supervised this work, and to K. M. Eisele and L. D. Gardner without whose cooperation the whole project of L-band balanced amplifiers would not have been successful.

APPENDIX

Theory of Symmetrical Directional Couplers

Since there is no literature readily available on the lumped constant directional coupler discussed in the text, this appendix is prepared to explain its principle from a slightly broader point of view. The theory to be presented is originally due to H. Seidel. It was developed during his association with Merrimac Research and Development, Inc., and is used in the design of their low frequency directional couplers.

For convenience, let us call two two-port networks "oppositely reflective" with respect to each other when they have identical scattering matrices except for opposite signs of their diagonal components. Now, let us consider the symmetrical network shown in Fig. 11, and apply incident waves of an even mode to ports 1 and 2, i.e., waves with the same amplitude and phase. Because of this symmetry, the actual (four-port) network can be considered as a two-port network acting on the incident mode. We thus have some reflection r_e of the even mode from ports 1 and 2, and a transmission t_e of the even mode to ports 3 and 4. Next, suppose that we apply incident waves of an odd mode to ports 1 and 2, i.e., waves with the same amplitude, but 180° out of phase. Again, because of the symmetry, the actual network acts as a two-port network to the incident mode and we have some reflection r_o and trans-

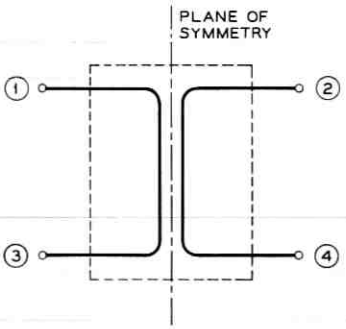


Fig. 11 — Symmetrical directional coupler with only one plane of symmetry.

mission t_e of the odd mode from ports 1 and 2 and from ports 3 and 4, respectively. With this much preparation, let us present the following theorem.

Theorem I: If the two two-port networks, presented by a symmetrical four-port network to its even and odd modes, are oppositely reflective, then the symmetrical four-port network is a directional coupler. (The converse is not necessarily true.)

The proof of this theorem is as follows. By definition, $r_e = -r_o$ and $t_e = t_o$. Therefore, let us define α and β by

$$\alpha = r_e = -r_o, \qquad \beta = t_e = t_o.$$

Suppose that a unit wave is incident at port 1. This can be considered as a superposition of even and odd modes incident at ports 1 and 2 with amplitudes of one half each. This fact can be expressed in matrix form

$$\begin{bmatrix} 1 \\ 0 \end{bmatrix} = \frac{1}{2}\begin{bmatrix} 1 \\ 1 \end{bmatrix} + \frac{1}{2}\begin{bmatrix} 1 \\ -1 \end{bmatrix}$$

where the upper and lower rows represent the waves on the left- and right-hand sides of the vertical plane of symmetry in Fig. 11, respectively. The reflection from ports 1 and 2 is therefore given by

$$\frac{r_e}{2}\begin{bmatrix} 1 \\ 1 \end{bmatrix} + \frac{r_o}{2}\begin{bmatrix} 1 \\ -1 \end{bmatrix} = \begin{bmatrix} 0 \\ \alpha \end{bmatrix}.$$

This means that port 1 has no reflection and port 2 has an output α. Similarly, the transmission to ports 3 and 4 is given by

$$\frac{t_e}{2}\begin{bmatrix}1\\1\end{bmatrix}+\frac{t_o}{2}\begin{bmatrix}1\\-1\end{bmatrix}=\begin{bmatrix}\beta\\0\end{bmatrix}.$$

This tells us that port 3 has an output β but no output appears at port 4. In other words, when a unit wave is incident to port 1, port 1 is matched, ports 2 and 3 have output waves α and β, respectively, and port 4 has no output. A similar argument holds for a unit wave incident at any other port of the symmetrical four-port network. This means that each port is matched and there is no coupling between conjugate ports. Thus, if a symmetrical four-port network is oppositely reflective to its even and odd modes in the sense discussed above, then it is a directional coupler and the theorem is proved.

The next theorem is useful for searching possible structures of directional couplers.

Theorem II: If two two-port networks with real generator and load immittances are dual in their normalized form with respect to the generator immittances, then the two-port networks are oppositely reflective independent of the frequency.

To make the meaning of the theorem clear and the proof easy, let us consider a simple example as shown in Fig. 12. For later use, the normalized load emittances are assumed to be unity in Fig. 12; however, this assumption is not necessary for the present discussion. The duality is satisfied when the normalized inductance l is equal to the normalized capacitance c. The theorem asserts that the networks inside the dotted lines are oppositely reflective at all frequencies. However, this is obvious from the following consideration. The normalized input (or output) impedance of one network is equal to the normalized input (or output) admittance of the other and therefore the reflection coefficients have equal magnitudes and opposite signs. The voltage across the load of one

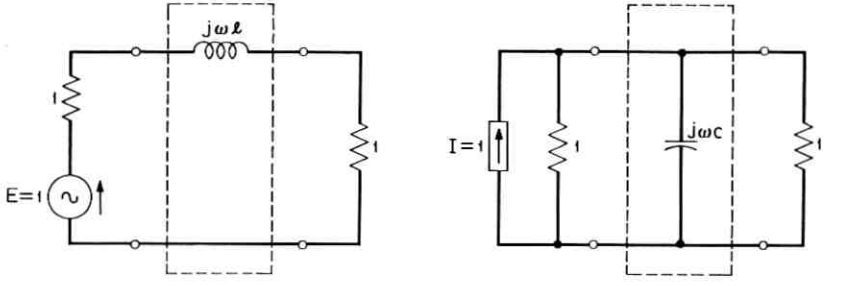


Fig. 12 — An example of dual circuits.

network is equal to the current flowing into the load of the other and therefore the transmissions are the same. Thus, regardless of the frequency, they are oppositely reflective with respect to each other. Since the above explanation is quite general, the theorem is proved.

Now, consider two closely spaced, thin and short parallel conductors and let us connect identical load Z_o's (real) at each end of the conductors. When the even mode is fed from one end of the conductors, the conductors should represent a lumped L. For the odd mode, the capacitance C between the conductors becomes effective while the currents in the two conductors cancel each other, giving no inductance effect. Therefore, from Theorems I and II this kind of circuit can work as a lumped constant directional coupler. In order to satisfy the required dual property, L/C has to be equal to Z_o^2. α and β can be calculated from Fig. 12 where the normalized inductance l corresponds to $2L/Z_o$. The coefficient 2 appears here because two loads Z_o are connected in parallel for the even mode to feed current to L.

If one realizes that coupled transmission lines exhibit a dual property to even and odd modes, the theory of the distributed coupler can be developed in a similar fashion. In the limiting case where the coupling factor k approaches unity and the electrical length θ of the coupled region approaches zero, the distributed coupler can be considered as a lumped constant directional coupler.

Although we have not discussed multisection directional couplers, they are useful in obtaining wider bandwidths. The following theorem serves as a guiding principle for constructing such couplers.

Theorem III: The oppositely reflective network of a cascade connection of two-port networks is equivalent to the cascade connection of the oppositely reflective two-port networks (provided that, for the comparison of opposite reflectivity, the same resistance is used for reference at each corresponding reference plane).

This theorem, together with Theorem I, guarantees that when several directional couplers of the type discussed above are connected in cascade, the resulting structure still works as a directional coupler. For the proof, let us first consider a cascade connection of two two-port networks. Using a signal flow graph similar to Fig. 10, the scattering matrix components of the cascade connection are given by

$$S_{11} = S_{11}(1) + \frac{S_{21}(1)S_{12}(1)S_{11}(2)}{1 - S_{22}(1)S_{11}(2)}$$

$$S_{12} = \frac{S_{12}(1)S_{12}(2)}{1 - S_{22}(1)S_{11}(2)}$$

$$S_{21}' = \frac{S_{21}(1)S_{21}(2)}{1 - S_{22}(1)S_{11}(2)}$$

$$S_{22} = S_{22}(2) + \frac{S_{22}(1)S_{21}(2)S_{12}(2)}{1 - S_{22}(1)S_{11}(2)}.$$

If we change the signs of $S_{11}(1)$, $S_{22}(1)$, $S_{11}(2)$, and $S_{22}(2)$ then the signs of S_{11} and S_{22} change but S_{12} and S_{21} remain the same. This means that the theorem is true for two networks in cascade. Next, let us increase the number of networks to 3, and first consider No. 1 network as one network and the cascade connection of No. 2 and No. 3 as the other. Then the application of the above proof for two networks shows that the opposite reflective network of the cascade connection of No. 1, No. 2 and No. 3 is equivalent to the cascade connection of the oppositely reflective network of No. 1 and that of No. 2 and No. 3 in cascade. Since another application of the above proof to the cascade connection of No. 2 and No. 3 shows that the oppositely reflective network of No. 2 and No. 3 in cascade is equivalent to the cascade connection of the oppositely reflective networks of No. 2 and No. 3, the theorem is proved for the case of three networks in cascade. Since this procedure of increasing the number of networks can be continued indefinitely, the proof of the theorem is completed.

REFERENCES

1. Engelbrecht, R. S., and Kurokawa, K., A Wideband Low Noise L-Band Balanced Transistor Amplifier, Proc. IEEE, *53*, No. 3, March, 1965, pp. 237–247.
2. Jones, E. M. T., and Bolljahn, J. T., Coupled Strip Transmission Line Filters and Directional Couplers, IRE Trans. MTT., *4*, April, 1956, pp. 75–81.
3. Kurokawa, K., Actual Noise Measure of Linear Amplifiers, Proc. IRE, *49*, Sept., 1961, pp. 1391–1397.
4. Experiments conducted by A. L. Stillwell and F. J. D'Alessio.

Digital Transmission in the Presence of Impulsive Noise*

By J. S. ENGEL

(Manuscript received April 4, 1965)

The transmission of digital data over telephone channels has been considered previously in the literature, and the effects of Gaussian noise have been analyzed. With experience, however, it has become apparent that while a background of Gaussian noise is present, the limiting noise is not Gaussian but impulsive in nature. It consists of bursts of high amplitude which occur at random considerably more often than is predicted by the rms value of the Gaussian background noise.

Measurements of the statistics of this noise have been initiated, and some results have been reported. In this paper, a model for the noise is constructed to be reasonably consistent with these measurements without becoming too complex to be handled analytically. Various modulation systems are analyzed to determine their performance in such a noise environment. Conditional error rates, in terms of the average number of bit errors per noise burst, are determined as functions of a convenient signal-to-noise ratio which is defined. The systems are ranked as to their performance in such a noise environment, and the ranking is found to be the same as that for Gaussian noise.

The improvement to be gained by employing complementary delay networks is investigated. Networks with linear and sinusoidal delay characteristics are considered.

I. INTRODUCTION

The limiting noise with regard to transmitting digital data is impulsive in nature. Consisting of bursts of high amplitude, it occurs at random considerably more often than is predicted by the rms value of the Gaussian background noise. In this paper, an analysis is made of digital transmission, by various modulation procedures, in such a noise environment.

* Taken from the dissertation submitted to the Faculty of the Polytechnic Institute of Brooklyn in partial fulfillment of the requirements for the degree of Doctor of Philosophy, 1964.

The block diagram of a generalized data transmission system including a telephone channel is shown in Fig. 1. The data signal consists of a train of either ideal impulses or square pulses, each pulse having either positive or negative polarity depending on whether it represents a mark or a space. The transmitter consists of a low-pass filter, which band limits the data signal, followed by a modulator and a band-pass filter. The low-pass filter is required to prevent "foldover" distortion in modulation. The band-pass filter restricts the transmitted signal to the range of frequencies passed by the channel. It avoids the waste of transmitted power in signal components which will not be received and also includes the channel splitting filters which prevent crosstalk into frequencies reserved for other channels. The low-pass and band-pass filters also shape the signal, giving it the form desired for transmission. The transmitted signal is applied to the channel, where it is contaminated by additive noise. The combined signal and noise enters a receiver which consists of a band-pass filter and demodulator which is generally followed by a low-pass filter and a synchronous decision device. The band-pass filter removes components of the noise outside the band of the signal, in addition to shaping the received signal. The low-pass filter, which is not required in every system, removes the demodulation products which lie outside the band of interest, and may also provide further signal shaping. The decision device samples the combined signal and noise at its input at discrete sampling instants, and on the basis of each sample, produces a mark or space symbol.

The channel is band limited, and may have an amplitude versus frequency characteristic which is not flat and a phase versus frequency characteristic which is not linear. However, for the purposes of analysis, the exact characteristics of the channel may be lumped with those of the receiver band-pass filter. The channel may then be considered as having the characteristics of an ideal band-pass filter, with unity gain and zero phase shift in the band of interest.

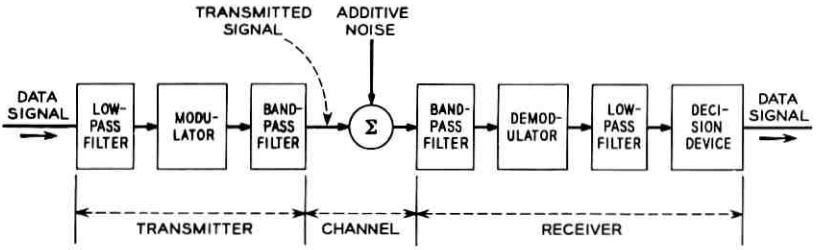


Fig. 1 — Generalized data transmission system.

The very broad objective, in considering such a problem, is to design the transmitter and receiver so that the number of errors caused by the noise is minimized, under the constraint that the average power of the transmitted signal is limited. A corollary objective, and obvious prerequisite, is to derive an expression for this error rate as a function of some readily measured signal-to-noise ratio, with the system characteristics as parameters. Then, one or two simple measurements of the noise on the channel will suffice to predict the error rate which can be expected of any system.

In this paper, various existing modulation systems are considered. The specific problem, then, is to find a suitable mathematical model for the noise, and for each system, to determine the response of the demodulation procedure to the combination of the signal and such noise. The processed noise at the output of the demodulator and filters is the source of the errors by the decision device. The optimization procedure consists of finding the decision criterion and filter characteristics at the receiver which minimize the number of errors caused by the noise, under the constraint of average power limiting of the transmitted signal. For each modulation system, this procedure specifies the receiver design, except for the phase characteristic of the receiving filter. In order to avoid intersymbol interference at the decision device, a specific functional form for the signal at that point is required. Given the filter characteristics at the receiver, as determined by the optimization procedure, the filter characteristics at the transmitter are made such that the response of the over-all transmission path to the applied data signal is the specific signal required at the decision device. This specifies the transmitter design for each modulation system.

In this paper, the following modulation systems are analyzed to determine their performance in the presence of impulsive noise:

(1.) Double sideband AM with coherent detection
(2.) Single sideband AM with coherent detection
(3.) AM with envelope detection
(4.) Frequency shift keying with frequency discrimination
(5.) Binary phase shift keying with differentially coherent detection
(6.) Quaternary phase shift keying with differentially coherent detection.

Conditional error rates, in terms of the average number of bit errors per noise burst, are determined as functions of a signal-to-noise ratio which is defined in the following section. The systems are ranked as to their performance, in order of increasing conditional error rate. This ranking is found to be the same as that for performance in the presence of Gaussian noise, and agrees with the ranking experienced in actual usage.

In the analyses, it is found that the phase characteristics of the receiving filter affect the performance of the system. This phenomenom, which is not present when the noise is Gaussian, leads to the consideration of complementary delay filters. Three types of delay filters are considered, and the improvement which results from their use is found, as a function of the maximum delay of the filter, for each of the modulation systems.

II. MODEL FOR THE IMPULSE NOISE

As previously described the various data transmission systems include modulation and demodulation of a carrier frequency. The data receivers each include a band-pass filter. For the purposes of analysis, the actual characteristics of the channel are lumped with those of the band-pass filter, and the channel is considered as having the characteristics of an ideal band-pass filter. With double-sideband transmission, these characteristics are made symmetrical about the carrier frequency. The impulse noise present on such a channel consists of bursts of carrier frequency ω_c, with random phase ψ, occurring randomly in time. When single-sideband transmission is used, the characteristics are asymmetrical and the Hilbert transform of the envelope is also present. This however, merely modifies the envelope and phase of the noise burst. The actual noise, as it occurs in the telephone plant, is wideband, with a spectrum covering many channels. The noise burst at the output of any one channel is the response of the channel to the wideband noise and contains only a small portion of the total spectrum. It is reasonable to assume therefore, that the spectrum of the noise burst at the output of the channel is essentially determined by the channel characteristics, with the original spectrum of the wideband noise having little effect. Under this assumption the envelope of each noise burst is the same, except for a random amplitude K. This is of course, an approximation. The spectra of the noise bursts do vary somewhat. However, in order to be useful, a model must be reasonably consistent with the observed phenomena without becoming too complex to be handled analytically. Approximating the envelopes of the noise burst by a representative time function $n_0(t)$ is a very reasonable approximation. At this point in the analysis, the envelope $n_0(t)$ is not limited to any particular function. A general expression for the error rate is derived, into which any specific function may be inserted. Then in order to obtain numerical results, a specific function is assumed. Up to that point, however, the analysis is quite general.

A single noise burst of carrier frequency $n(t)$, occurring at time $t = \tau$, may be represented in the form:

$$n(t) = Kn_0(t - \tau) \cos(\omega_c t + \psi) \tag{1}$$

where K, τ, and ψ are three independent random variables associated with each burst. For ease of notation, the envelope $n_0(t)$ is normalized to have an energy of two watt-seconds. Then, the energy ε of the noise burst $n(t)$ is equal to K^2.

Measurements have been made on representative telephone channels, to determine the statistical characteristics of the noise.[1,3] As a result, there have been functional forms for:

$$F(x) = Pr[K > x] \tag{2}$$

suggested by Fennick,[3] Mertz,[5,6] and others, which fit the measured data reasonably well for the range of K large enough to cause errors. These are families of functions $F_\alpha(x)$, where the parameter α may vary from channel to channel but is constant for any one channel. One such family of functions, suggested by Mertz, which fits the data reported on by Fennick, is computationally suited to this development. A reference amplitude K_0 is chosen sufficiently low in value that any noise burst which might cause an error would have an amplitude greater than K_0. The relative frequencies of amplitudes greater than K_0 are measured and plotted. That is to say, the conditional probability

$$\Pr[K > x \mid K > K_0]$$

is plotted versus the relative amplitude

$$20 \log_{10}(x/K_0) db.$$

These plots are straight lines on logarithmic paper, as shown in Fig. 2. The negative slope of each line, in decades per 10 db, is the parameter α. These plots satisfy the equation

$$\Pr[K > x \mid K > K_0] = (K_0/x)^{2\alpha}. \tag{3}$$

When describing the impulse noise, it is sufficient to consider only those bursts with amplitude greater than K_0; if K_0 is chosen low enough that bursts with smaller amplitude do not cause errors, these smaller bursts may be ignored and artificially assumed not to occur. Then, the family of functions F_α(a) is given by:

$$F_\alpha(x) = \Pr[K > x] = (K_0/x)^{2\alpha} \quad \text{for} \quad x > K_0. \tag{4}$$

The majority of the measured distributions have values of α between 1

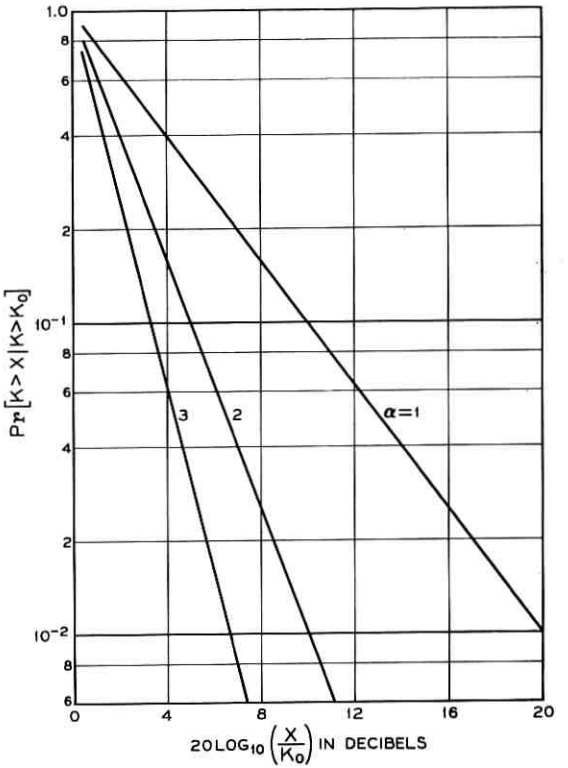


Fig. 2 — Distribution of noise burst amplitudes.

and 2.5, with occasional higher values. High values of α correspond to channels with a small percentage of high amplitude bursts, and thus result in low conditional error rates. Lower values of α correspond to channels with a greater percentage of high amplitude bursts, resulting in higher conditional error rates. A distribution with α equal to 1 would have an infinite variance; the mean of K^2, which is the average energy in a noise burst, would be infinite. A value of 1 is a lower bound on α, resulting in an upper bound on the conditional error rate.

A function which is more useful in the ensuing analysis is the distribution of the energies of the noise bursts. Since the envelope function is normalized such that this energy ε is equal to K^2, letting $y = x^2$ and $\varepsilon_0 = K_0^2$ yields:

$$Q(y) = \Pr[\varepsilon > y] = (\varepsilon_0/y)^\alpha \quad \text{for} \quad y > \varepsilon_0 \tag{5}$$

where ε_0 is the energy of the lowest energy noise burst included in the distribution.

The result of the ensuing analysis is the calculation of the conditional error rate, in terms of the average number of errors per noise burst, as a function of a signal-to-noise ratio. This ratio is defined as the average signal energy per data bit divided by the minimum burst energy ε_0. For a given modulation system and signal power, a value of ε_0 may be chosen that is sufficiently low to insure that all noise bursts which cause errors must have greater energy. Then for any given channel, an impulse counter with its threshold set to count impulses with energy greater than ε_0 can find the number of such noise bursts per data bit transmitted. The product of this number and the conditional error rate is then the over-all error rate, in bits in error per bit transmitted.

The second random variable associated with each noise burst is the phase ψ of the carrier, which is assumed to be uniformly distributed in the interval between 0 and 2π.

The third random variable associated with each noise burst is the time of occurrence. Let the sampling instant for an arbitrarily chosen data pulse be designated as the time origin $t = 0$, and let the time of occurrence of the closest noise burst, past or future, be designated as $t = \tau$. It is assumed that the average interval between noise bursts is so much greater than the duration of the bursts, that the probability of two of them overlapping is negligibly small. Then, if an error does occur due to noise at the sampling instant $t = 0$, only the nearest burst, occurring at $t = \tau$, has contributed toward it. For the average rates at which the noise bursts have been observed to occur, this assumption is valid.

For the purposes of the subsequent analyses, the density function $p(\tau)$ is only of interest for absolute values of τ less than the duration of the noise burst. If the closest noise burst occurs at a time τ which is more than a noise burst duration removed from the sampling instant, then no remnant of the noise is present at the sampling instant, and hence no error can occur. This duration is so short compared to the average interval between bursts that $p(\tau)$ is essentially constant, equal to $p(0)$, over that range of values. Further, for the distributions which have been measured, $p(0)$ is approximately equal to the average number of bursts, per unit time, β, of greater energy than ε_0. (This can be shown to be exactly true for the case when the intervals are exponentially distributed, with a Poisson distribution for the number of noise bursts occurring in a given period. It is very nearly true for the other distributions proposed.) For absolute values of τ greater than the duration of the noise burst,

the density function $p(\tau)$ does not enter the analysis, and no further assumptions are required.

In the following sections, the various modulation systems are analyzed, and their performance in the presence of the noise described above is determined.

III. AMPLITUDE MODULATION SYSTEMS

3.1 *Double-Sideband AM with Coherent Detection*

The ideal data receiver using coherent detection consists of a band-pass filter $H_c(\omega)$, followed by a coherent detector and a post-detection low pass filter $H_L(\omega)$. When double sideband transmission is utilized, the band-pass filter is symmetrical about the carrier frequency ω_c. The detector multiplies the received signal by a carrier signal with the same frequency and phase as that of the modulated carrier at the transmitter. The post-detection filter removes those components of the resulting signal centered about $| \omega | = 2\omega_c$, in addition to shaping the resulting baseband signal. The post-detection filter is followed by a synchronous decision device which samples its input at discrete sampling instants T seconds apart. A block diagram of the receiver is shown in Fig. 3.

At the input to the decision device, the data signal $y(t)$ consists of a sequence of identical, except for polarity, pulses $y_0(t)$ spaced T seconds apart

$$y(t) = \sum_i a_i E y_0(t - iT) \tag{6}$$

where a_i equals $+1$ if the ith bit is a mark, and -1 if it is a space. At the jth sampling instant $t = jT$, the data signal $y(jT)$ should equal a_jE with no intersymbol interference present from any other pulse. The signal $y_0(t)$ is band limited and cannot be restricted to a time slot T seconds in width. Each pulse extends over several sampling instants. In order to avoid intersymbol interference, $y_0(t)$ is constrained to be a

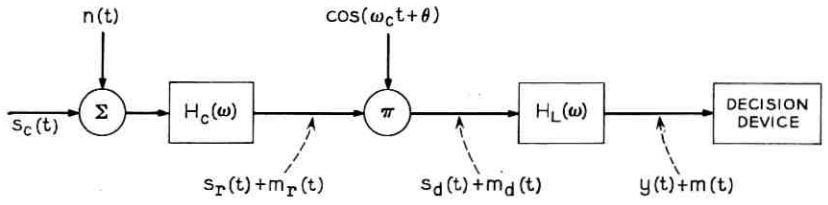


Fig. 3 — AM receiver with coherent detection.

member of a class of functions equal to 1 at $t = 0$ and equal to zero at all other integral multiples of T, satisfying Nyquist's First Criterion.[8]

In the absence of noise, each sample may have one of only two possible values. The ith sample may be $\pm E$, depending on whether the ith bit is a mark or a space. When noise is present, the sample may have one of a range of values. The decision device is a maximum likelihood estimator, producing a mark symbol if the sample value is positive, and a space symbol if it is negative. At any sampling instant, the noise will cause an error if it has an absolute value greater than E and polarity opposite to that of the data signal. The probability of this event occurring, as a function of the signal-to-noise ratio, will now be found.

Let $t = 0$ denote the sampling instant for an arbitrarily chosen data pulse, and let $t = \tau$ denote the time of occurrence of the closest noise burst. As discussed, the probability of two bursts overlapping is negligibly small. If an error does occur at the sampling instant, it is the result only of the closest noise burst. As described in Section II, the noise burst $n(t)$ is given by (1), and has a spectrum given by

$$N(\omega) = (K/2)\{e^{j\psi}N_0(\omega - \omega_c) \exp [-j(\omega - \omega_c)\tau] + e^{-j\psi}N_0(\omega + \omega_c) \exp [-j(\omega + \omega_c)\tau]\}. \tag{7}$$

The spectrum of the noise at the output of the post-detection filter $H_L(\omega)$ is

$$M(\omega) = (K/4)N_0(\omega)[e^{j(\psi-\theta)}H_c(\omega + \omega_c) + e^{-j(\psi-\theta)}H_c(\omega - \omega_c)]e^{-j\omega\tau}H_L(\omega). \tag{8}$$

As described, the band-pass filter $H_c(\omega)$ is symmetrical about the carrier frequency $|\omega| = \omega_c$, therefore,

$$H_c(\omega + \omega_c)H_L(\omega) = H_c(\omega - \omega_c)H_L(\omega).$$

An "equivalent receiving filter" $H(\omega)$ is now defined as

$$H(\omega) \triangleq H_c(\omega + \omega_c)H_L(\omega) = H_c(\omega - \omega_c)H_L(\omega), \tag{9}$$

so that

$$M(\omega) = (K/2) \cos \phi N_0(\omega)H(\omega) \exp (-j\omega\tau) \tag{10}$$

where $\phi = \psi - \theta$ is uniformly distributed in the interval between 0 and 2π. Let $m_0(t)$ be defined as the response of the equivalent receiving filter $H(\omega)$ to the normalized envelope $n_0(t)$, given by

$$m_0(t) = \frac{1}{2\pi}\int_{-\infty}^{\infty} N_0(\omega)H(\omega) \exp (j\omega t)\, d\omega. \tag{11}$$

Then, at the output of the post-detection filter, the noise has the spectrum

$$M(\omega) = (K/2) \cos \phi \, M_0(\omega) \exp(-j\omega\tau). \tag{12}$$

At the sampling instant $t = 0$, the noise at the input to the decision device is equal to

$$m(0) = (K/2) \cos \phi m_0(-\tau). \tag{13}$$

An error will occur if the absolute value of $m(0)$ is greater than E and if the polarity of the noise is opposite to that of the signal. The probability of error at the sampling instant $t = 0$, given that the closest noise burst has occurred at $t = \tau$ with relative phase ϕ, is thus equal to

$$\Pr[\text{error} \mid \tau,\phi] = \tfrac{1}{2} \Pr[m^2(0) > E^2 \mid \tau,\phi]. \tag{14}$$

Substituting the right-hand side of (13) for $m(0)$, a rearrangement of terms yields

$$\Pr[\text{error} \mid \tau,\phi] = \tfrac{1}{2} \Pr\left[\varepsilon > \frac{4E^2}{\cos^2 \phi m_0^2(-\tau)} \mid \tau,\phi\right]. \tag{15}$$

By (5), this is

$$\Pr[\text{error} \mid \tau,\phi] = \frac{1}{2}\left[\frac{\varepsilon_0 \cos^2 \phi m_0^2(-\tau)}{4E^2}\right]^{\alpha}. \tag{16}$$

The relative phase ϕ is uniformly distributed in the interval between 0 and 2π. Therefore, the probability of error at the sampling instant $t = 0$, given that the closest noise burst has occurred at $t = \tau$, is given by

$$\Pr[\text{error} \mid \tau] = \frac{C_\alpha}{2}\left[\frac{\varepsilon_0 m_0^2(-\tau)}{4E^2}\right]^{\alpha}, \tag{17}$$

where

$$C_\alpha = \frac{1}{2\pi}\int_0^{2\pi} (\cos^2 \phi)^\alpha \, d\phi. \tag{18}$$

The probability of error, at any arbitrary sampling instant, is

$$\Pr[\text{error}] = \frac{C_\alpha}{2}\left[\frac{\varepsilon_0}{4E^2}\right]^{\alpha} \int_{-\infty}^{\infty} m_0^{2\alpha}(-\tau) p(\tau) \, d\tau. \tag{19}$$

The density function $p(\tau)$ is only of interest for the range of τ less than the duration of the noise burst $m_0(t)$. For larger values of τ, $m_0^{2\alpha}(-\tau)$ is essentially equal to zero, and the integrand in (19) is zero. As de-

scribed in the discussion of the noise model, the density function is essentially constant and equal to the average number of bursts per unit time β, for the range of τ of interest. The probability of error is therefore essentially equal to

$$\Pr[\text{error}] = \frac{\beta C_\alpha}{2}\left[\frac{\varepsilon_0}{4E^2}\right]^\alpha \int_{-\infty}^{\infty} m_0^{2\alpha}(-\tau)\, d\tau. \tag{20}$$

The system is constrained to operate with a limitation on the average power of the signal on the channel. Under such a limitation, the error rate should be given as a function of the average power of the transmitted signal, rather than of the peak value E. It may be shown that for the double-sideband amplitude modulated signal $s_c(t)$, which yields the signal $y(t)$ given by (6), the average power on the channel is

$$\bar{W} = \frac{E^2}{\pi T}\int_{-\infty}^{\infty}\left|\frac{Y_0(\omega)}{H(\omega)}\right|^2 d\omega \tag{21}$$

where $Y_0(\omega)$ is the spectrum of the individual data pulse $y_0(t)$. Let I_1 be defined as the integral

$$I_1 \triangleq \frac{1}{2\pi T}\int_{-\infty}^{\infty}\left|\frac{Y_0(\omega)}{H(\omega)}\right|^2 d\omega. \tag{22}$$

Then substituting in (21) and (22) into (20),

$$\Pr[\text{error}] = \frac{\beta C_\alpha}{2}\left[\frac{\varepsilon_0 I_1}{2\bar{W}}\right]^\alpha \int_{-\infty}^{\infty} m_0^{2\alpha}(-\tau)\, d\tau. \tag{23}$$

Substituting $t = -\tau$ in the definite integral and rearranging some of the terms, (23) yields

$$\Pr[\text{error}] = \frac{\beta C_\alpha T}{2}\left[\frac{1}{2} I_1 \frac{\varepsilon_0}{\bar{W}T}\right]^\alpha \int_{-\infty}^{\infty} [T m_0^2(t)]^\alpha \frac{dt}{T}, \tag{24}$$

where $\bar{W}T/\varepsilon_0$ is the signal-to-noise ratio defined previously as the signal energy per bit divided by the energy in the minimum noise burst. The rearrangement of terms in the above expression normalizes the integral to a dimensionless constant.

Since $Y_0(\omega)$ is specified by the system designer, and $N_0(\omega)$ is known, the equivalent receiving filter characteristic $H(\omega)$ may be optimized in the sense that the probability of error is minimized. For each value of α, however, a different characteristic is optimum. Since the data system is to be used over randomly selected channels, with all possible values of α, it appears reasonable that it should be optimized in a minimax sense.

The system design should therefore be optimized for the worst case, which corresponds to a value of α equal to 1. If the system then operates with an acceptably low error rate over a channel with a value of α equal to 1, it will operate with an even lower error rate over another channel with a higher value of α. The optimization procedure therefore consists of finding the filter characteristic which minimizes the probability of error for a value of α equal to one. For that case, the probability of error is equal to

$$\Pr[\text{error}] = \frac{\beta}{16\pi}\frac{\varepsilon_0}{\bar{W}T}\int_{-\infty}^{\infty}\left|\frac{Y_0(\omega)}{H(\omega)}\right|^2 d\omega \int_{-\infty}^{\infty} m_0^2(t)\, dt \tag{25}$$

where, by Parseval's theorem

$$\begin{aligned}\int_{-\infty}^{\infty} m_0^2(t)\, dt &= \frac{1}{2\pi}\int_{-\infty}^{\infty} |\, M_0(\omega)\,|^2\, d\omega \\ \int_{-\infty}^{\infty} m_0^2(t)\, dt &= \frac{1}{2\pi}\int_{-\infty}^{\infty} |\, N_0(\omega)H(\omega)\,|^2\, d\omega.\end{aligned} \tag{26}$$

Then

$$\Pr[\text{error}] = \frac{1}{32\pi^2}\left[\frac{\varepsilon_0}{\bar{W}T}\right]\int_{-\infty}^{\infty}\left|\frac{Y_0(\omega)}{H(\omega)}\right|^2 d\omega \int_{-\infty}^{\infty} |\, N_0(\omega)H(\omega)\,|^2\, d\omega. \tag{27}$$

By Schwarz's inequality

$$\left|\int_{-\infty}^{\infty} Y_0(\omega)N_0(\omega)\, d\omega\right|^2 \leqq \int_{-\infty}^{\infty}\left|\frac{Y_0(\omega)}{H(\omega)}\right|^2 d\omega \int_{-\infty}^{\infty} |\, N_0(\omega)H(\omega)\,|^2\, d\omega. \tag{28}$$

The equality is satisfied when

$$\frac{Y_0(\omega)}{H(\omega)} = CN_0^*(\omega)H^*(\omega) \tag{29}$$

where C is any real constant and the asterisk denotes the complex conjugate. Since the left hand side of the inequality is independent of $H(\omega)$, it represents the minimum value the right hand side may take. The probability of error is therefore minimized when the equality is satisfied, and this occurs when the filter characteristic satisfies the relation

$$|\, H(\omega)\,|^2 = \frac{1}{C}\frac{Y_0(\omega)}{N_0^*(\omega)}. \tag{30}$$

The phase characteristic of $H(\omega)$ has no effect on the probability of error for the limiting case when α is equal to 1. The system performance is therefore evaluated for zero (or linear) phase shift. The phase char-

acteristic does affect the error probability when α is greater than 1; this effect is considered in detail in a subsequent section on complementary delay filters.

The probability of error is dependent on two parameters of the noise bursts, the distribution of their amplitudes-defined by α, and their relative frequency — defined by β. This second dependence consists merely of a direct proportionality, and need not be carried along in the computation. The performance measure may be normalized by considering the conditional error rate, defined as the average number of bit errors per noise burst. The probability of error is equal to the average number of bit errors per bit transmitted. If this number is divided by the average number of noise bursts per bit transmitted, βT, the result is the average number of bit errors per noise burst,

$$\bar{N} = \frac{\Pr[\text{error}]}{\beta T}.$$

given by:

$$\bar{N} = \frac{C_\alpha}{2}\left[\frac{1}{2} I_1 \frac{\varepsilon_0}{\bar{W}T}\right]^\alpha \int_{-\infty}^{\infty} [T m_0^2(t)]^\alpha \frac{dt}{T}. \tag{31}$$

In order to obtain some numerical results, the general expression for $\bar{N}$ given above is to be evaluated for a special case which is of interest. In the transmission of data over narrow-band telephone channels, the spectrum of the individual data pulses $y_0(t)$ is often the "raised cosine" spectrum given by

$$Y_0(\omega) = (T/2)[1 + \cos(\omega T/2)] \quad \text{for} \quad |\omega| < 2\pi/T. \tag{32}$$

This signal satisfies all three of Nyquist's criteria:

(1.) $y_0(iT) = \begin{cases} 1 \text{ for } i = 0 \\ 0 \text{ for } i = 1, 2, 3, \cdots. \end{cases}$

(2.) $y_0[i(T/2)] = \begin{cases} \frac{1}{2} \text{ for } i = 1 \\ 0 \text{ for } i = 2, 3, 4, \cdots. \end{cases}$

(3.) The envelope of $y_0(t)$ approaches zero very rapidly as $|t|$ increases.

For the sake of simplicity it will be assumed that in the narrow band of interest $|\omega| < 2\pi/T$, the spectrum of the noise $N_0(\omega)$ is constant. Since the energy of the noise burst $n_0(t)$ in the band of interest is equal to 2 watt-seconds, the noise spectrum is given by

$$N_0(\omega) = \sqrt{T} \quad \text{for} \quad |\omega| < 2\pi/T. \tag{33}$$

The optimum filter characteristic is given by

$$\begin{aligned} H(\omega) &= \{\tfrac{1}{2}[1 + \cos(\omega T/2)]\}^{\frac{1}{2}} \quad \text{for} \quad |\omega| < 2\pi/T \\ H(\omega) &= \cos(\omega T/4) \qquad\qquad\qquad\quad \text{for} \quad |\omega| < 2\pi/T. \end{aligned} \tag{34}$$

Note that the optimization permitted an arbitrary scale factor for $H(\omega)$, since signal and noise would be identically affected. This scale factor has been chosen such that the maximum value of $H(\omega)$ is unity.

The expression for $\bar{N}$ has been numerically evaluated on a digital computer, for values of α equal to 1, 2, and 3, chosen to be representative. The results are tabulated here, and are plotted in Fig. 4.

α	$\bar{N}$
1	$0.124\ \varepsilon_0/\bar{W}T$
2	$0.0545\ (\varepsilon_0/\bar{W}T)^2$
3	$0.0302\ (\varepsilon_0/\bar{W}T)^3$

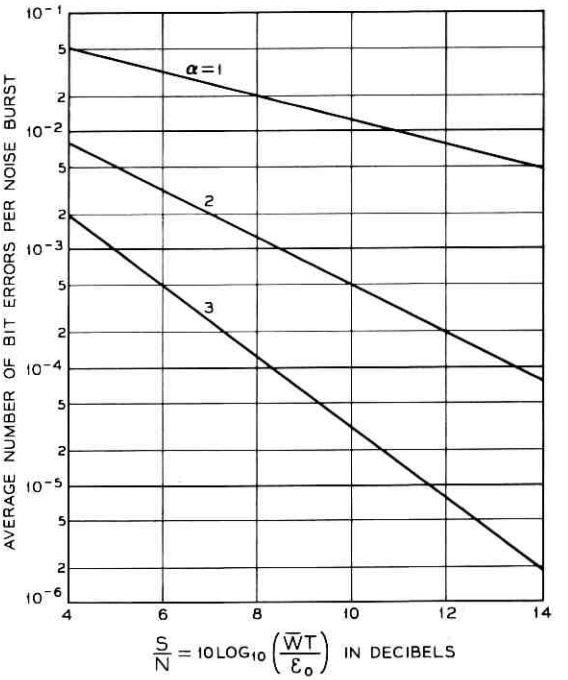


Fig. 4 — Performance of double-sideband AM system with coherent detection.

3.2 *Single-Sideband AM with Coherent Detection*

The ideal data receiver has the same form as the one described in the previous section. The only difference is that when single-sideband transmission is utilized, the band-pass filter $H_c(\omega)$ does not pass frequencies below ω_c

$$H_c(\omega) = 0 \quad \text{for} \quad |\omega| < \omega_c . \tag{35}$$

At frequencies above ω_c, the band-pass filter is identical to that for the double-sideband case.

As in the previous analysis, let $t = 0$ denote the sampling instant for an arbitrarily chosen data pulse, and let $t = \tau$ denote the time of occurrence of the closest noise burst. As in the double-sideband case, the spectrum of the noise burst at the output of the post-detection filter is given by (8). However, since $H_c(\omega)$ is equal to zero for all $|\omega|$ less than ω_c, this spectrum is equal to

$$M(\omega) = (K/4)M_0(\omega)\,[\cos\phi + jsgn(\omega)\sin\phi]\exp(-j\omega\tau), \tag{36}$$

and the noise burst is

$$m(t) = (K/4)[m_0(t-\tau)\cos\phi - \hat{m}_0(t-\tau)\sin\phi]$$

where $\hat{m}_0(t)$ is the Hilbert transform of $m_0(t)$, given by:

$$\hat{m}_0(t) = \frac{1}{2\pi}\int_{-\infty}^{\infty}\frac{1}{j}\,sgn(\omega)N_0(\omega)H(\omega)\exp(j\omega t)\,d\omega. \tag{37}$$

At the sampling instant $t = 0$, the noise at the input to the decision device is equal to

$$m(0) = (K/4)[m_0(-\tau)\cos\phi - \hat{m}_0(-\tau)\sin\phi]. \tag{38}$$

Substituting the right-hand side of (38) for $m(0)$ in the expression for the probability of error, (14), and rearranging terms,

$$\Pr[\text{error} \mid \tau,\phi] = \tfrac{1}{2}\Pr\left\{\varepsilon > \frac{16E^2}{[m_0(-\tau)\cos\phi - \hat{m}_0(-\tau)\sin\phi]^2}\,\middle|\,\tau,\phi\right\}. \tag{39}$$

By (5), this is equal to

$$\Pr[\text{error} \mid \tau,\phi] = \frac{1}{2}\left\{\frac{\varepsilon_0[m_0(-\tau)\cos\phi - \hat{m}_0(-\tau)\sin\phi]^2}{16E^2}\right\}^{\alpha}, \tag{40}$$

and may be rewritten in the form

$$\Pr[\text{error} \mid \tau,\phi] = \frac{1}{2}\left[\frac{\varepsilon_0}{16E^2}\right]^{\alpha} [m_0^2(-\tau) + \hat{m}_0^2(-\tau)]^{\alpha} (\cos^2 \chi)^{\alpha} \quad (41)$$

where, for any given value of τ, the angle

$$\chi = \phi + \tan^{-1} \frac{\hat{m}_0(-\tau)}{m_0(-\tau)} \quad (42)$$

is uniformly distributed in the interval between 0 and 2π. Averaging (41) with respect to χ and τ yields

$$\Pr[\text{error}] = \frac{\beta C_\alpha}{2}\left[\frac{\varepsilon_0}{16E^2}\right]^{\alpha} \int_{-\infty}^{\infty} [m_0^2(-\tau) + \hat{m}_0^2(-\tau)]^{\alpha}\, d\tau. \quad (43)$$

It may be shown that, for the single-sideband amplitude-modulated signal $s_c(t)$ which yields the signal $y(t)$ given by (6), the average power on the channel is equal to

$$\bar{W} = 4E^2 I_1 \quad (44)$$

where I_1 is the integral defined by (22). Substituting (44) in (43) and rearranging some of the terms yields:

$$\Pr[\text{error}] = \frac{\beta C_\alpha T}{2}\left[\frac{1}{4} I_1 \frac{\varepsilon_0}{\bar{W}T}\right]^{\alpha} \int_{-\infty}^{\infty} [T m_0^2(t) + T\hat{m}_0^2(t)]^{\alpha} \frac{dt}{T} \quad (45)$$

where $\bar{W}T/\varepsilon_0$ is the signal-to-noise ratio defined previously.

In the same manner as for the double-sideband system, the equivalent receiving filter characteristic $H(\omega)$ is designed to minimize the probability of error for the case when α is equal to 1. For that value of α, the error probability is given by

$$\Pr[\text{error}] = \frac{\beta}{32\pi}\left[\frac{\varepsilon_0}{\bar{W}T}\right] \int_{-\infty}^{\infty} \left|\frac{Y_0(\omega)}{H(\omega)}\right|^2 d\omega \left[\int_{-\infty}^{\infty} m_0^2(t)\, dt + \int_{-\infty}^{\infty} \hat{m}_0^2(t)\, dt\right]. \quad (46)$$

By Parseval's theorem

$$\int_{-\infty}^{\infty} m_0^2(t)dt = \int_{-\infty}^{\infty} \hat{m}_0^2(t)dt = \frac{1}{2\pi}\int_{-\infty}^{\infty} |N_0(\omega)H(\omega)|^2\, d\omega. \quad (47)$$

Substituting (47) into (46) yields the same expression as for the double-sideband case, given by (27). The optimum filter characteristic is therefore the same, satisfying (30).

The conditional error rate, defined as the average number of bit errors per noise burst, is given by

$$\bar{N} = \frac{C_\alpha}{2}\left[\frac{1}{4} I_1 \frac{\varepsilon_0}{\bar{W}T}\right]^\alpha \int_{-\infty}^{\infty} [Tm_0^2(t) + T\hat{m}_0^2(t)]^\alpha \frac{dt}{T}. \quad (48)$$

In order to obtain numerical results, the general expression for $\bar{N}$ is evaluated for the special case described previously. The spectrum of an individual data pulse is the raised cosine spectrum, and the noise spectrum is assumed to be constant in the band of interest. The expression for $\bar{N}$ has been numerically evaluated for values of α equal to 1, 2 and 3. The results are tabulated here, and are plotted in Fig. 5.

α	$\bar{N}$
1	$0.124\ (\varepsilon_0/\bar{W}T)$
2	$0.0241\ (\varepsilon_0/\bar{W}T)^2$
3	$0.0069\ (\varepsilon_0/\bar{W}T)^3$

3.3 *AM with Envelope Detection*

The ideal data receiver using envelope detection consists of a receiving filter $H_c(\omega)$, symmetrical about the carrier frequency ω_c, followed by

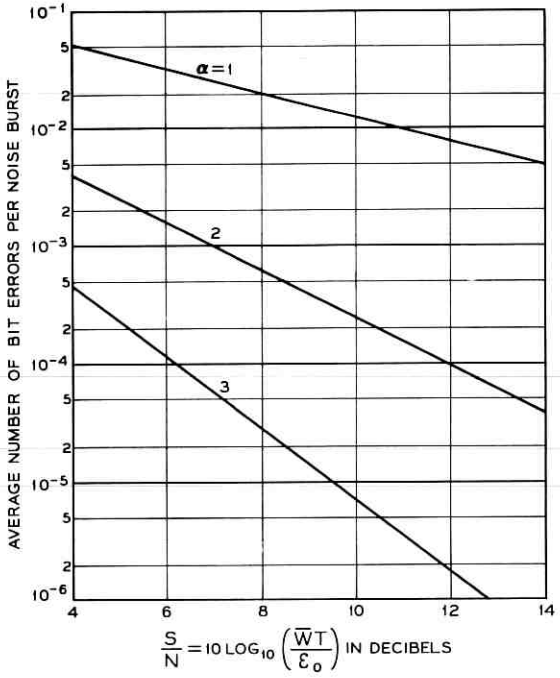


Fig. 5 — Performance of single-sideband AM system with coherent detection.

an ideal envelope detector. The output of the envelope detector is applied to a synchronous decision device. A block diagram of the receiver is shown in Fig. 6. The signal $y(t)$ at the input to the decision device has the form of (6). At the ith sampling instant, it is equal to

$$y(iT) = |a_i|.$$

Since polarity information is lost, and only magnitude information is delivered to the decision device, unipolar signaling must be used. If the ith bit is a mark, a_i is equal to $+1$; if the ith bit is a space, a_i is equal to zero. The threshold of the decision logic is set at a level λE, where λ has a value between zero and one. If the ith sample is greater than λE, a mark symbol is produced; if less than λE, a space symbol is produced. The value of λ is chosen to minimize the probability of error in the presence of noise.

Let $t = 0$ denote the sampling instant for an arbitrarily chosen data pulse, and let $t = \tau$ denote the time of occurrence of the closest noise burst. The noise burst on the channel is given by (1). At the input to the detector, the noise burst is given by

$$m_r(t) = Km_0(t - \tau)[\cos\phi \cos(\omega_c t + \theta) - \sin\phi \sin(\omega_c t + \theta)] \quad (49)$$

where the previously described phase difference

$$\phi = \psi - \theta,$$

is uniformly distributed in the interval between 0 and 2π.

At the input to the detector, the combined signal and noise voltage is equal to

$$\begin{aligned} v_i(t) = {} & [y_r(t) + Km_0(t - \tau)\cos\phi]\cos(\omega_c t + \theta) \\ & - Km_0(t - \tau)\sin\phi \sin(\omega_c t + \theta). \end{aligned} \quad (50)$$

The detector output is

$$v_0(t) = \sqrt{[y_r(t) + Km_0(t - \tau)\cos\phi]^2 + [Km_0(t - \tau)\sin\phi]^2}. \quad (51)$$

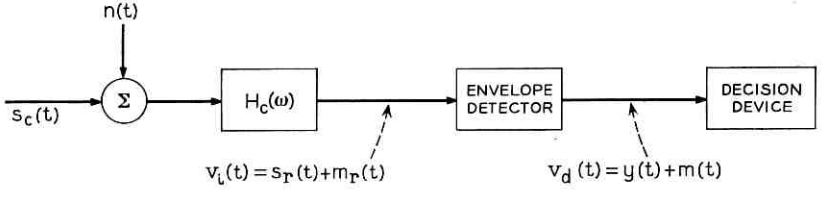


Fig. 6 — AM receiver with envelope detection.

If, at the sampling instant $t = 0$, a space has been sent, the voltage at the input to the decision device is equal to the noise alone

$$v_0(0) = |Km_0(-\tau)|. \tag{52}$$

An error will occur if this magnitude is greater than λE. The probability of error at the sampling instant $t = 0$, given that a space has been sent and that the closest noise burst has occurred at $t = \tau$, is therefore equal to

$$\Pr[\text{error} \mid \text{space}, \tau] = \Pr[\varepsilon m_0^2(-\tau) > \lambda^2 E^2 \mid \tau] \tag{53}$$

and, by (5), this is

$$\Pr[\text{error} \mid \text{space}, \tau] = \left[\frac{\varepsilon_0 m_0^2(-\tau)}{\lambda^2 E^2}\right]^\alpha \tag{54}$$

If, at the sampling instant $t = 0$, a mark has been sent, the voltage at the input to the decision device is equal to

$$v_0(0) = +\sqrt{[E + Km_0(-\tau)\cos\phi]^2 + [Km_0(-\tau)\sin\phi]^2} \tag{55}$$

and an error will occur if this is less than λE. This is shown graphically in the phasor diagrams of Fig. 7. An error will occur if the vector sum of the signal E and the noise $Km_0(-\tau)$ making an angle ϕ with E falls within the circle of radius λE.

Fig. 7(a) shows the phasor diagram when $m_0(-\tau)$ is positive. An error can occur only for values of ϕ in the interval

$$\pi + \sin^{-1}\lambda \geqq \phi \geqq \pi - \sin^{-1}\lambda.$$

For all other values of ϕ the vector sum cannot fall within the circle. When ϕ is within the above interval, an error will occur if $Km_0(-\tau)$

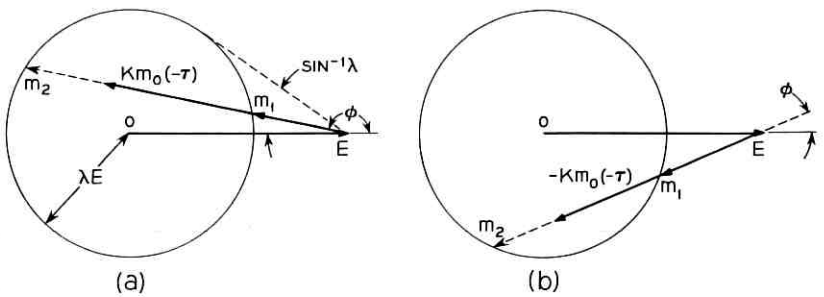


Fig. 7 — (a) Phasor diagram when $m_0(-\tau)$ is positive. (b) Phasor diagram when $m_0(-\tau)$ is negative.

lies in the interval

$$m_2 > Km_0(-\tau) > m_1 \tag{56}$$

where m_1 and m_2 are functions of ϕ. By the Law of Cosines, m_1 and m_2 must satisfy

$$\lambda^2 E^2 = E^2 + m^2 - 2mE \cos(\pi - \phi), \tag{57}$$

hence,

$$\begin{aligned} m_1 &= \frac{(1-\lambda^2)E}{\cos(\pi-\phi) + \sqrt{\lambda^2 - \sin^2(\pi-\phi)}} \\ m_2 &= \frac{(1-\lambda^2)E}{\cos(\pi-\phi) - \sqrt{\lambda^2 - \sin^2(\pi-\phi)}} \end{aligned} \tag{58}$$

The probability of error at the sampling instant $t = 0$, given that a mark has been sent and that the closest noise burst has occurred at $t = \tau$, such that $m_0(-\tau) > 0$, with phase difference ϕ, is equal to the probability that the noise is within the range (56), and is

$$\begin{aligned} &\Pr[\text{error} \mid \text{mark}, \tau, m_0(-\tau) > 0, \phi] \\ &\quad = \Pr\left[\frac{m_2^2}{m_0^2(-\tau)} > \varepsilon > \frac{m_1^2}{m_0^2(-\tau)} \,\middle|\, \tau, \phi\right] \\ &\qquad \text{for} \quad \pi + \sin^{-1}\lambda > \phi > \pi - \sin^{-1}\lambda \end{aligned} \tag{59}$$

and is equal to zero for all other values of ϕ. By (5) this is equal to

$$\begin{aligned} &\Pr[\text{error} \mid \text{mark}, \tau, m_0(-\tau) > 0, \phi] \\ &= \left\{\frac{\varepsilon_0 m_0^2(-\tau)[\cos(\pi-\phi) + \sqrt{\lambda^2 - \sin^2(\pi-\phi)}]^2}{(1-\lambda^2)^2E^2}\right\}^\alpha \\ &\quad - \left\{\frac{\varepsilon_0 m_0^2(-\tau)[\cos(\pi-\phi) - \sqrt{\lambda^2 - \sin^2(\pi-\phi)}]^2}{(1-\lambda^2)^2E^2}\right\}^\alpha \\ &\qquad \text{for} \quad \pi + \sin^{-1}\lambda > \phi > \pi - \sin^{-1}\lambda. \end{aligned} \tag{60}$$

The phase difference ϕ is uniformly distributed in the interval between 0 and 2π. Therefore, averaging (60) with respect to ϕ yields:

$$\begin{aligned} &\Pr[\text{error} \mid \text{mark}, \tau, m_0(-\tau) > 0] \\ &= \frac{1}{\pi}\left[\frac{\varepsilon_0 m_0^2(-\tau)}{(1-\lambda^2)^2E^2}\right]^\alpha \int_0^{\sin^{-1}\lambda} \{[\cos\phi + \sqrt{\lambda^2 - \sin^2\phi}]^{2\alpha} \\ &\quad - [\cos\phi - \sqrt{\lambda^2 - \sin^2\phi}]^{2\alpha}\}\, d\phi. \end{aligned} \tag{61}$$

Fig. 7(b) shows the phasor diagram when $m_0(-\tau)$ is negative. By reasoning similar to that presented above, it can be shown that the probability of error at the sampling instant $t = 0$, given that a mark has been sent and that the closest noise burst has occurred at $t = \tau$, such that $m_0(-\tau) < 0$, is also given by (61). The probability of error given that $m_0(-\tau)$ is positive is identically equal to the probability of error given that it is negative. Since the two events are disjoint

$$\Pr[\text{error} \mid \text{mark}, \tau] = \frac{1}{\pi}\left[\frac{\varepsilon_0 m_0^2(-\tau)}{(1-\lambda^2)^2 E^2}\right]^{\alpha} \int_0^{\sin^{-1}\lambda} \{[\cos\phi + \sqrt{\lambda^2 - \sin^2\phi}]^{2\alpha} - [\cos\phi - \sqrt{\lambda^2 - \sin^2\phi}]^{2\alpha}\}\, d\phi. \tag{62}$$

Marks and spaces are assumed to occur with equal probability. Therefore, the probability of error at the sampling instant $t = 0$, given that the closest noise burst has occurred at $t = \tau$, is equal to one-half the sum of (54) and (62). Let a signal-to-noise ratio coefficient be defined by

$$f_\alpha(\lambda) = \left\{\frac{1}{\lambda^{2\alpha}} + \frac{1}{\pi(1-\lambda^2)^{2\alpha}} \int_0^{\sin^{-1}\lambda} \{[\cos\phi + \sqrt{\lambda^2 - \sin^2\phi}]^{2\alpha} - [\cos\phi - \sqrt{\lambda^2 - \sin^2\phi}]^{2\alpha}\}\, d\phi\right\}^{1/\alpha}. \tag{63}$$

Then, the probability of error, given τ, equals

$$\Pr[\text{error} \mid \tau] = \frac{1}{2}\left[\frac{f_\alpha(\lambda)\varepsilon_0 m_0^2(-\tau)}{E^2}\right]^{\alpha}. \tag{64}$$

As shown previously, averaging with respect to τ yields that the probability of error at any arbitrary sampling instant is essentially equal to:

$$\Pr[\text{error}] = \frac{\beta}{2}\left[\frac{f_\alpha(\lambda)\varepsilon_0}{E^2}\right]^{\alpha} \int_{-\infty}^{\infty} m_0^{2\alpha}(-\tau)\, d\tau. \tag{65}$$

The value of the threshold level λ is to be selected to minimize the coefficient $f_\alpha(\lambda)$. The functions $f_\alpha(\lambda)$ and $df_\alpha(\lambda)/d\lambda$ have been evaluated for values of α equal to 1, 2, and 3. These are plotted in Figs. 8 and 9, respectively. Since a value of α equal to 1 yields the highest error rate, the threshold level λ should be selected to minimize $f_1(\lambda)$. The error rates for other values of α will always be lower than that for a value of α equal to 1. However, $f_1(\lambda)$ exhibits the broadest minimum and is most amenable to compromise. As may be seen in Fig. 8, a threshold level λ

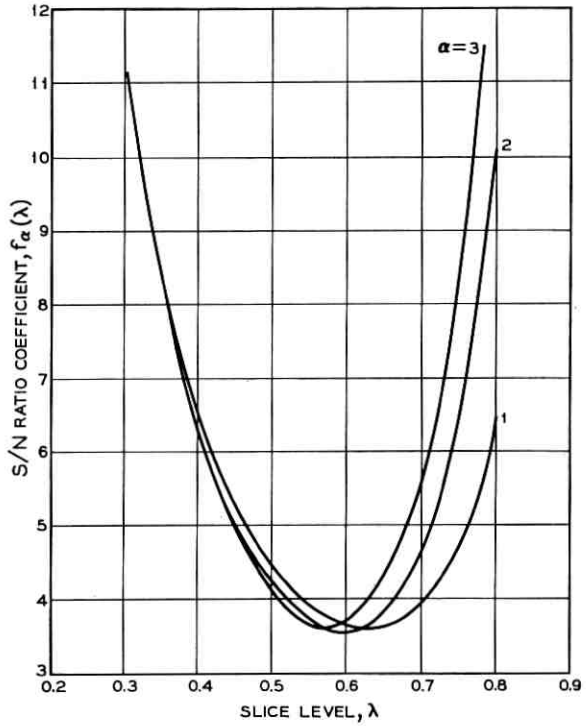


Fig. 8 — Signal-to-noise ratio coefficient.

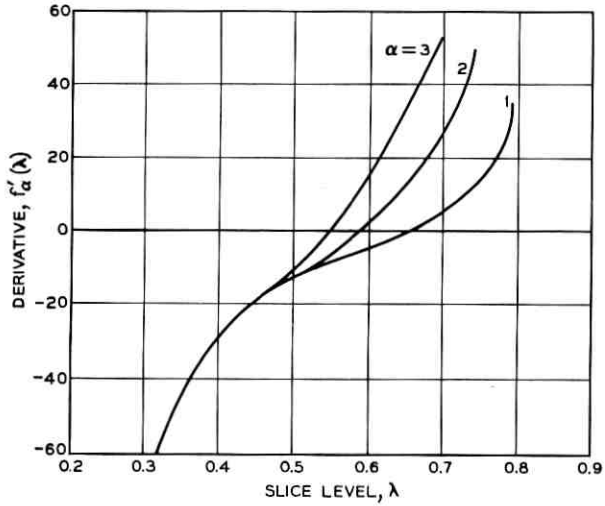


Fig. 9 — Rate of change of signal-to-noise ratio coefficient.

equal to 0.6 is a satisfactory compromise. At that value of λ, f_1 equals 3.65, f_2 equals 3.55, and f_3 equals 3.65.

It may be shown that for the unipolarly keyed modulated signal $s_c(t)$, the average power on the channel is equal to

$$\bar{W} = E^2\left[\frac{1}{8}I_1 + \frac{1}{8T^2}\left|\frac{Y_0(0)}{H(0)}\right|^2\right]. \tag{66}$$

The second term is recognized as the result of a carrier frequency component which is present because the average value of the baseband data signal is not equal to zero. Substituting (66) into (65) and rearranging some of the terms,

$$\Pr[\text{error}] = \frac{\beta T}{2}\left\{\frac{f_\alpha \varepsilon_0\left[I_1 + \frac{1}{T^2}\left|\frac{Y_0(0)}{H(0)}\right|^2\right]}{8\bar{W}T}\right\}^\alpha \int_{-\infty}^{\infty}[Tm_0^2(t)]^\alpha \frac{dt}{T} \tag{67}$$

where $\bar{W}T/\varepsilon_0$ is the signal-to-noise ratio defined previously.

In the same manner as for the previously considered systems, the filter characteristic $H(\omega)$ is selected to minimize the error probability when α is equal to one. For that value of α, the probability of error is given by

$$\Pr[\text{error}] = \frac{\beta f_1}{(8\pi)^2}\left[\frac{\varepsilon_0}{\bar{W}T}\right]\int_{-\infty}^{\infty}\left|\frac{Y_0(\omega)}{H(\omega)}\right|^2\left[1 + \frac{2\pi}{T}\delta(\omega)\right]d\omega \cdot \int_{-\infty}^{\infty}|N_0(\omega)H(\omega)|^2\,d\omega. \tag{68}$$

By Schwarz's inequality, the above expression is minimized when the amplitude characteristic of $H(\omega)$ satisfies the relation

$$|H(\omega)|^2 \sim \frac{Y_0(\omega)}{N_0^*(\omega)}\left[1 + \frac{2\pi}{T}\delta(\omega)\right]^{\frac{1}{2}}. \tag{69}$$

The requirement of an impulse at zero frequency is equivalent to carrier suppression at the transmitter. The infinite gain at the receiver then restores the dc component of the base-band data signal. Since the construction of such a filter is not feasible, it is assumed that the filter $H(\omega)$ has the suboptimum "smooth" characteristic given by (30).

The conditional error rate, defined as the average number of bit errors per noise burst, is

$$\bar{N} = \frac{1}{2}\left\{\frac{1}{8}f_\alpha\left[I_1 + \frac{1}{T^2}\left|\frac{Y_0(0)}{H(0)}\right|^2\right]\frac{\varepsilon_0}{\bar{W}T}\right\}^\alpha \int_{-\infty}^{\infty}[Tm_0^2(t)]^\alpha \frac{dt}{T}. \tag{70}$$

In order to obtain numerical results, the general expression for $\bar{N}$ is

evaluated for the special case described previously. The spectrum of an individual data pulse is the raised cosine spectrum, and the spectrum of the noise is assumed to be constant in the band of interest. The expression for $\bar{N}$ has been numerically evaluated on a digital computer, for values of α equal to 1, 2 and 3. The results are tabulated below, and are plotted in Fig. 10.

α	$\bar{N}$
1	$0.455\ (\varepsilon_0/\bar{W}T)$
2	$0.456\ (\varepsilon_0/\bar{W}T)^2$
3	$0.586\ (\varepsilon_0/\bar{W}T)^3$

IV. FREQUENCY SHIFT KEYING SYSTEM

The ideal data receiver for a frequency shift keying system consists of a receiving filter $H_c(\omega)$ which is symmetrical about the carrier frequency ω_c, followed by an ideal frequency discriminator. The output of

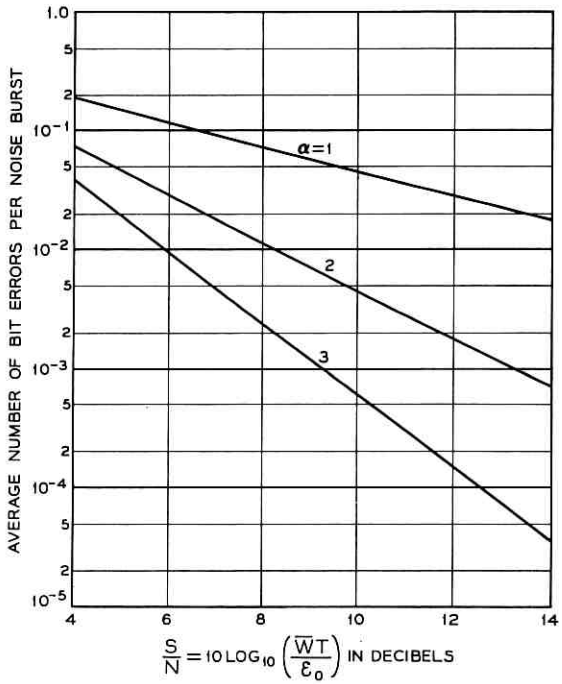


Fig. 10 — Performance of AM system with envelope detection.

the discriminator is the instantaneous frequency ω_i of the signal (plus noise) at its input. The discriminator is followed by a synchronous decision device identical to the one in the coherent AM systems. A block diagram of the receiver is shown in Fig. 11.

Sunde[9] has shown that in the absence of noise, intersymbol interference can be eliminated even though the signal at the input to the detector must be bandlimited. The frequency shift signal is generated by mixing the outputs of two synchronized oscillators; one oscillating at a mark frequency ω_m, the other at a space frequency ω_s. For this discussion, let the mark frequency be higher than the space frequency; the choice is arbitrary. The carrier frequency is defined as the mid-frequency

$$\omega_c = (\omega_m + \omega_s)/2$$

and the shift frequency is defined by:

$$2\omega_d = \omega_m - \omega_s .$$

The modulated signal is

$$s(t) = \tfrac{1}{2}\{[1 + g(t)] \cos [(\omega_c + \omega_d)t + \theta] - [1 - g(t)][\cos (\omega_c - \omega_d)t + \theta]. \quad (71)$$

At the sampling instants, the mixing function is equal to

$$g(iT) = \pm 1$$

so that the instantaneous frequency at the sampling instants is

$$\omega_i(iT) = \omega_c \pm \omega_d$$

depending on whether the ith bit is a mark or a space.

The modulated signal may be rewritten as

$$s(t) = \sin \omega_d t \sin (\omega_c t + \theta) - g(t) \cos \omega_d t \cos (\omega_c t + \theta). \quad (72)$$

After successive filtering, amplification, and transmission, the signal at the input to the discriminator, in the absence of noise, is equal to

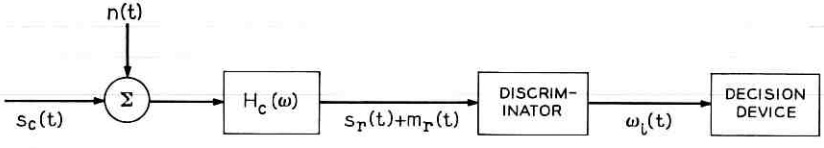


Fig. 11 — Receiver for frequency shift keyed system.

$$s_r(t) = P(t) \cos (\omega_c t + \theta) - Q(t) \sin (\omega_c t + \theta), \tag{73}$$

where

$$\begin{aligned} P(t) &= -y(t) \\ Q(t) &= -E \sin \omega_d t. \end{aligned} \tag{74}$$

As a further restriction, the shift frequency is made equal to the bit rate, so that

$$\omega_d = \pi/T$$

and

$$y(t) = \sum_i (-1)^i a_i E y_0(t - iT), \tag{75}$$

where, as before, a_i is $+1$ if the ith bit is a mark and -1 if it is a space, and $y_0(t)$ satisfies Nyquist's First Criterion. Sunde has shown that these conditions yield a signal which in the absence of noise presents no intersymbol interference.

As in the analysis of the AM systems, let $t = 0$ denote an arbitrarily selected data sampling instant, and let $t = \tau$ denote the time of occurrence of the closest noise burst. The noise at the output of the receiving filter is given by (49). With both signal and noise present at the discriminator input, the output is

$$\omega_i(t) = \frac{d}{dt}\left[\tan^{-1} \frac{Q(t) + y(t)}{P(t) + x(t)}\right] \tag{76}$$

where

$$\begin{aligned} x(t) &= Km_0(t - \tau) \cos \phi \\ y(t) &= Km_0(t - \tau) \sin \phi \end{aligned} \tag{77}$$

and is equal to

$$\omega_i(t) = \frac{[P(t) + x(t)][\dot{Q}(t) + \dot{y}(t)] - [Q(t) + y(t)][\dot{P}(t) + \dot{x}(t)]}{[P(t) + x(t)]^2 + [Q(t) + y(t)]^2}. \tag{78}$$

Since the denominator of $\omega_i(t)$ is always positive, the decision device will produce a mark if

$$\begin{aligned} V = [-Ea_0 + Km_0(-\tau) \cos \phi][-(E\pi/T) + K\dot{m}_0(-\tau) \sin \phi] \\ - [Km_0(-\tau) \sin \phi][\dot{P}(0) + K\dot{m}_0(-\tau) \cos \phi] \end{aligned} \tag{79}$$

is positive, and a space if it is negative.

For the cases of interest, it can be shown that $\dot{P}(0)$ is considerably less than $\dot{x}(0)$ for those noise bursts which cause errors. In addition, $\dot{P}(0)$ is equal to zero. For these reasons, the $\dot{P}(0)$ term is dropped in the subsequent steps, since to retain it would unnecessarily complicate the analysis.

When a mark has been sent, and a_0 is equal to $+1$, V is negative and an error occurs when

$$K[(T/\pi)\dot{m}_0(-\tau)\sin\phi + m_0(-\tau)\cos\phi] > E. \tag{80}$$

The probability of error at the sampling instant $t = 0$, given that a mark has been sent and that the closest noise burst has occurred at $t = \tau$ with phase difference ϕ, is equal to:

$$\Pr[\text{error} \mid \text{mark},\tau,\phi] = \frac{1}{2}\Pr\left\{\varepsilon^2 > \frac{E^2}{\left[\frac{T}{\pi}\dot{m}_0(-\tau)\sin\phi + m_0(-\tau)\cos\phi\right]^2}\,\middle|\,\tau,\phi\right\} \tag{81}$$

and by (5), this is equal to

$$\Pr[\text{error} \mid \text{mark},\tau,\phi] = \frac{1}{2}\left\{\frac{\varepsilon_0\left[\frac{T}{\pi}\dot{m}_0(-\tau)\sin\phi + m_0(-\tau)\cos\phi\right]^2}{E^2}\right\}^{\alpha}. \tag{82}$$

When a space has been sent, and a_0 is equal to -1, V is positive and an error occurs when:

$$K[(T/\pi)\dot{m}_0(-\tau)\sin\phi - m_0(-\tau)\cos\phi] > E. \tag{83}$$

Following the reasoning presented above, the probability of error at the sampling instant $t = 0$, given that a space has been sent and that the closest noise burst has occurred at $t = \tau$ with phase difference ϕ, is equal to

$$\Pr[\text{error} \mid \text{space},\tau,\phi] = \frac{1}{2}\left\{\frac{\varepsilon_0\left[\frac{T}{\pi}\dot{m}_0(-\tau)\sin\phi - m_0(-\tau)\cos\phi\right]^2}{E^2}\right\}^{\alpha}. \tag{84}$$

Marks and spaces occur with equal probability. Therefore, the probability of error, given τ and ϕ, is

$$\Pr[\text{error} \mid \tau,\phi] = \frac{1}{4}\left[\frac{\varepsilon_0}{E^2}\right]^{\alpha}\left\{\left[\frac{T}{\pi}\dot{m}_0(-\tau)\right]^2 + [m_0(-\tau)]^2\right\}^{\alpha} \cdot \left\{\left\{\cos^2\left[\phi + \tan^{-1}\frac{T}{\pi}\frac{\dot{m}_0(-\tau)}{m_0(-\tau)}\right]\right\}^{\alpha} + \left\{\cos^2\left[\phi - \tan^{-1}\frac{T}{\pi}\frac{\dot{m}_0(-\tau)}{m_0(-\tau)}\right]\right\}^{\alpha}\right\}. \tag{85}$$

Averaging with respect to ϕ and τ yields

$$\Pr[\text{error}] = \frac{\beta C_\alpha}{2}\left[\frac{\varepsilon_0}{E_2}\right]^{\alpha}\int_{-\infty}^{\infty}\left\{\left[\frac{T}{\pi}\dot{m}_0(-\tau)\right]^2 + [m_0(-\tau)]^2\right\}^{\alpha} d\tau. \tag{86}$$

It can be shown that, for the frequency shift keyed signal $s_c(t)$, the average power on the channel is

$$\bar{W} = E^2[(1/4A_0^2) + \tfrac{1}{2}I_1], \tag{87}$$

where

$$A_0 = |H(\pm\omega_s)| = |H(\pm\omega_m)|. \tag{88}$$

The first term in the bracket of (87) is the result of discrete components of the signal at the mark and space frequencies. Substituting (87) into (86) and rearranging some of the terms yields

$$\Pr[\text{error}] = \frac{\beta C_\alpha T}{2}\left[\frac{\varepsilon_0\left(\frac{1}{2A_0^2} + I_1\right)}{2\bar{W}T}\right]^{\alpha} \cdot \int_{-\infty}^{\infty}\left\{T\left[\frac{T}{\pi}\dot{m}_0(t)\right]^2 + T[m_0(t)]^2\right\}^{\alpha}\frac{dt}{T} \tag{89}$$

where $\bar{W}T/\varepsilon_0$ is the signal-to-noise ratio defined previously.

In the same manner as for the systems considered previously, an equivalent receiving filter characteristic $H(\omega)$ is to be found which minimizes the error rate when α is equal to 1. For that value of α, the probability of error is equal to

$$\Pr[\text{error}] = \frac{\beta T}{8}\left[\frac{\varepsilon_0}{\bar{W}T}\right]\left[\frac{1}{2A_0^2} + \frac{1}{2\pi T}\int_{-\infty}^{\infty}\left|\frac{Y_0(\omega)}{H(\omega)}\right|^2 d\omega\right] \cdot \int_{-\infty}^{\infty}\left[\frac{T^2}{\pi^2}\dot{m}_0^2(t) + m_0^2(t)\right]dt, \tag{90}$$

where

$$\frac{1}{A_0^2} = \frac{1}{2}\int_{-\infty}^{\infty}\left[\delta\left(\omega - \frac{\pi}{T}\right) + \delta\left(\omega + \frac{\pi}{T}\right)\right]\left|\frac{1}{H(\omega)}\right|^2 d\omega. \tag{91}$$

By Parseval's theorem

$$\int_{-\infty}^{\infty} \dot{m}_0^2(t)\, dt = \frac{1}{2\pi} \int_{-\infty}^{\infty} \omega^2 \mid M_0(\omega) \mid^2 d\omega.$$

Therefore, for a value of α equal to 1, (90) may be rewritten

$$\begin{aligned} \Pr[\text{error}] = {} & \frac{\beta}{(4\pi)^2} \left[\frac{\varepsilon_0}{\bar{W}T}\right] \int_{-\infty}^{\infty} \left\{ \frac{\pi T}{2} \left[\delta\left(\omega - \frac{\pi}{T}\right) + \delta\left(\omega + \frac{\pi}{T}\right) \right] \right. \\ & \left. + \mid Y_0(\omega) \mid^2 \right\} \left| \frac{1}{H(\omega)} \right|^2 d\omega \cdot \int_{-\infty}^{\infty} \left(\frac{T^2}{\pi^2} \omega^2 + 1 \right) \mid N_0(\omega) H(\omega) \mid^2 d\omega. \end{aligned} \tag{92}$$

For ease of notation, let

$$\begin{aligned} a(\omega) &\triangleq \pi T/2\{\delta[\omega - (\pi/T)] + \delta[\omega + (\pi/T)]\} + \mid Y_0(\omega) \mid^2 \\ b(\omega) &\triangleq [(T^2/\pi^2)\omega^2 + 1] \mid N_0(\omega) \mid^2 \\ h(\omega) &\triangleq \mid H(\omega) \mid^2. \end{aligned} \tag{93}$$

To minimize the probability of error, it is necessary to find the function $h(\omega)$ which minimizes the product.

$$P = \int_{-\infty}^{\infty} \frac{a(\omega)}{h(\omega)}\, d\omega \int_{-\infty}^{\infty} b(\omega) h(\omega)\, d\omega. \tag{94}$$

The minimum occurs when the variation

$$\begin{aligned} \delta P = {} & \int_{-\infty}^{\infty} \frac{a(\omega)}{h(\omega)}\, d\omega \int_{-\infty}^{\infty} b(\omega)\, \delta h(\omega)\, d\omega \\ & - \int_{-\infty}^{\infty} \frac{a(\omega)}{h^2(\omega)}\, \delta h(\omega_2)\, d\omega \int_{-\infty}^{\infty} b(\omega) h(\omega)\, d\omega \end{aligned} \tag{95}$$

$$\delta P = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \left[\frac{a(\omega_1)}{h(\omega_1)}\, b(\omega_2) - \frac{a(\omega_2)}{h^2(\omega_2)}\, b(\omega_1) h(\omega_1) \right] \delta h(\omega_2)\, d\omega_1\, d\omega_2$$

is equal to zero. Since it must be equal to zero for any variation $\delta h(\omega_2)$, the bracketed term must be zero

$$\frac{a(\omega_1)}{h(\omega_1)}\, b(\omega_2) - \frac{a(\omega_2)}{h^2(\omega_2)}\, b(\omega_1) h(\omega_1) = 0 \tag{96}$$

for all values of ω_1 and ω_2. This requires that $h^2(\omega) b(\omega)/a(\omega)$ be equal to a constant. It can be shown that the second variation $\delta^2 P$ is positive when the above condition is met, so that P is truly at a minimum, and not at a maximum or stationary point. To minimize the probability of error, the filter characteristic must satisfy

$$| H(\omega) |^2 \sim \left\{ \frac{\frac{\pi T}{2}\left[\delta\left(\omega - \frac{\pi}{T}\right) + \delta\left(\omega + \frac{\pi}{T}\right)\right] + | Y_0(\omega) |^2}{\left(\frac{T^2}{\pi^2}\omega^2 + 1\right) | N_0(\omega) |^2} \right\}^{\frac{1}{2}}. \quad (97)$$

The impulses in the receiving filter characteristics correspond to suppression of the discrete components at the transmitting filter. Since it is not feasible to construct such a filter, it is assumed that the suboptimum "smooth" filter, given by

$$| H(\omega) |^2 \sim \frac{| Y_0(\omega) |}{\left(\frac{T^2}{\pi^2}\omega^2 + 1\right)^{\frac{1}{2}} | N_0(\omega) |} \quad (98)$$

is to be used.

The conditional error rate, defined as the average number of bit errors per noise burst, is given by

$$\bar{N} = \frac{C_\alpha}{2}\left[\frac{\varepsilon_0}{\bar{W}T}\right]^\alpha \left[\frac{1}{4A_0} + \frac{1}{2} I_1\right]^\alpha \cdot \int_{-\infty}^{\infty} \left\{ T\left[\frac{T}{\pi}\dot{m}_0(t)\right]^2 + T[m_0(t)]^2 \right\}^\alpha \frac{dt}{T}. \quad (99)$$

The general expression for $\bar{N}$ is evaluated for the same special case described previously. The spectrum of an individual data pulse is the raised cosine spectrum, and the spectrum of the noise is assumed to be constant in the band of interest. The expression for $\bar{N}$ has been numerically evaluated, on a digital computer, for values of α equal to 1, 2, and 3. The results are tabulated below, and are plotted in Fig. 12.

α	$\bar{N}$
1	$0.402\ \varepsilon_0/\bar{W}T$
2	$0.392\ (\varepsilon_0/\bar{W}T)^2$
3	$0.48\ (\varepsilon_0/\bar{W}T)^3$

V. PHASE SHIFT KEYING SYSTEM WITH DIFFERENTIALLY COHERENT DETECTION

5.1 *Binary (Two Phase) System*

The ideal data receiver utilizing differentially coherent detection consists of a receiving filter $H_c(\omega)$ symmetrical about the carrier frequency ω_c, followed by a detector. The differentially coherent detector multi-

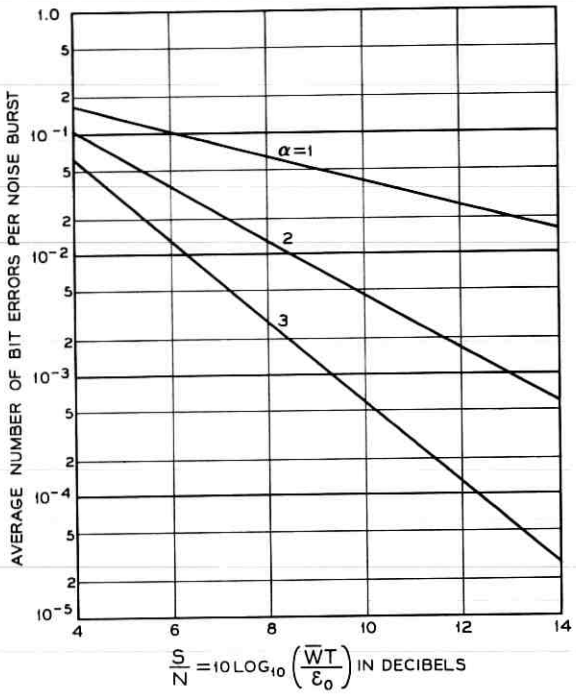


Fig. 12 — Performance of frequency shift keyed system.

plies the received signal by the signal which had been received one bit duration earlier. The output of the detector is applied to a synchronous decision device of the type described previously, in the analysis of the double-sideband AM system. A block diagram of the receiver is shown below, in Fig. 13.

In the absence of noise, the signal at the input to the detector is of the form

$$s_r(t) = y(t) \cos \omega_c t \tag{100}$$

where the modulating signal $y(t)$ satisfies (6) The multiplier a_i may be ± 1. If the ith bit is a mark, a_i is made equal to a_{i-1} ; if a space, a_i equal to $-a_{i-1}$. In the absence of noise, the signal at the output of the detector is equal to

$$v_0(t) = y(t)y(t - T) \cos \omega_c(t) \cos \omega_c(t - T). \tag{101}$$

The carrier frequency ω_c is selected to be an integral multiple of the bit

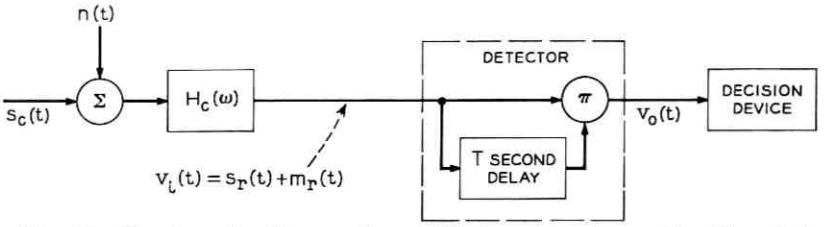


Fig. 13 — Receiver for binary phase shift keyed system with differentially coherent detection.

rate, so that $\omega_c T$ is an integral multiple of 2π. At the ith sampling instant, the signal at the output of the detector is equal to

$$v_0(iT) = y(iT)y[(i-1)T] = a_i a_{i-1} E^2. \tag{102}$$

The sample is equal to $+E^2$ if the ith bit is a mark, and $-E^2$ if it is a space. The decision device produces a mark symbol if the sample is positive and a space symbol if it is negative.

Let $t = 0$ denote the sampling instant for an arbitrarily chosen data pulse, and let $t = \tau$ denote the time of occurrence of the closest noise burst. At the sampling instant $t = 0$, the output of the detector is

$$v_0(0) = [a_0 E + K m_0(-\tau)\cos\psi][a_{-1}E + K m_0(-\tau - T)\cos\psi]. \tag{103}$$

To find the probability of error, given τ and ψ, it is necessary to find the ranges of values of K which cause the polarity of $v_0(0)$ to be reversed. For ease of notation, let the following two functions of τ and ψ be defined:

$$\begin{aligned} B_1(\tau,\psi) &\triangleq \frac{E}{m_0(-\tau)\cos\psi} \\ B_2(\tau,\psi) &\triangleq \frac{E}{m_0(-\tau - T)\cos\psi}. \end{aligned} \tag{104}$$

The output of the detector, at the sampling instant $t = 0$, may be written

$$v_0(0) = E^2[a_0 + (K/B_1)][a_{-1} + (K/B_2)]. \tag{105}$$

Depending on the values of τ and ψ, B_1 and B_2 may each be either positive or negative, so that there are four possible combinations of polarity. In addition, either B_1 or B_2 may have the larger absolute value. Each of these eight possible combinations of polarity and relative absolute value must be investigated separately. The range of K which causes a reversal of polarity, depending on the values of a_0 and a_{-1}, is to be found. This is most clearly done graphically. In Fig. 14, the eight possible combina-

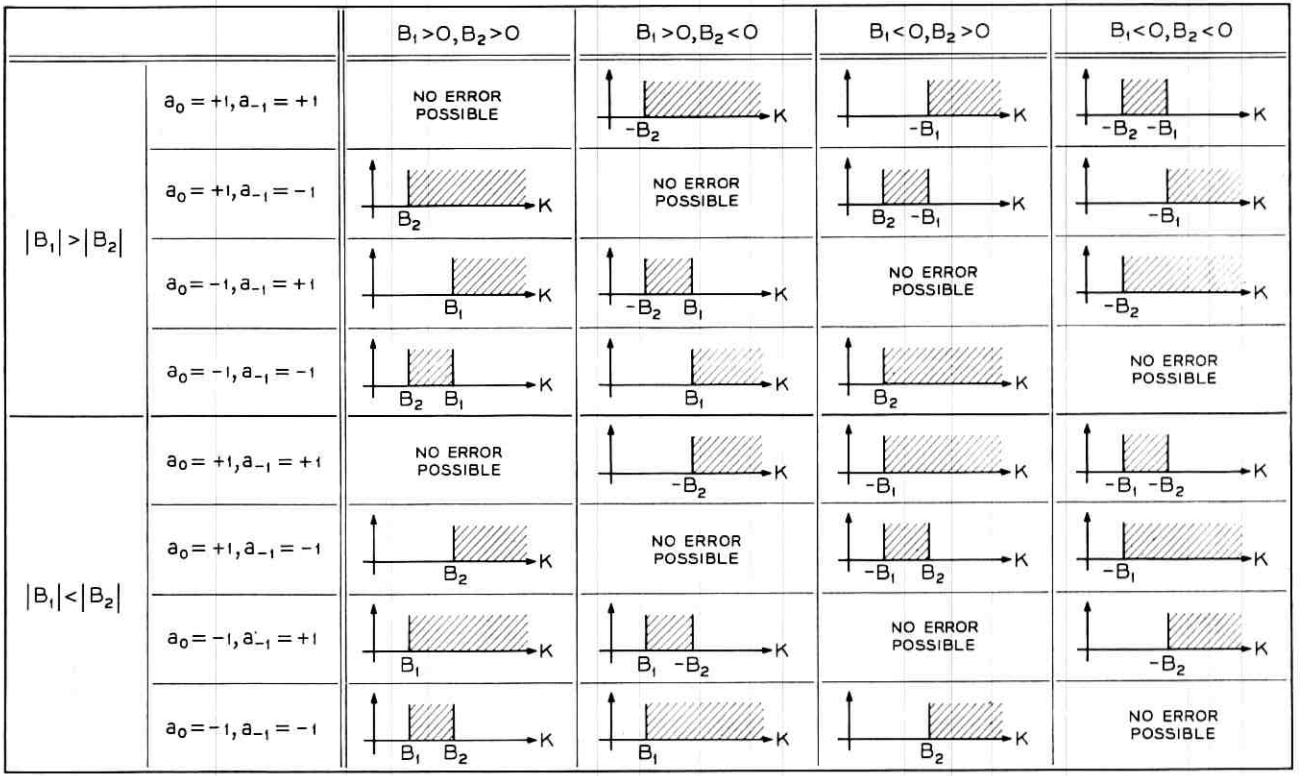


Fig. 14 — Error regions for the noise burst amplitude.

tions are shown separately and, for each of the four equally probable combinations of a_0 and a_{-1}, the range of K which causes an error is shown shaded. Because the four combinations of a_0 and a_{-1} are equally probable, each with probability of occurrence equal to $\frac{1}{4}$, the probability of error, given τ and ψ, for any one of the eight possibilities for B_1 and B_2, is equal to

$$\Pr[\text{error} \mid \tau,\psi] = \tfrac{1}{4} \sum_{a_0,a_{-1}} \Pr[K \text{ in shaded region}]. \tag{106}$$

It may be seen in Fig. 14 that, for each of the eight possibilities, this is equal to

$$\Pr[\text{error} \mid \tau,\psi] = \tfrac{1}{2} \Pr[K > \min (\mid B_1 \mid, \mid B_2 \mid) \mid \tau,\psi] \tag{107}$$

The probability of error, given τ and ψ, may be rewritten,

$$\Pr[\text{error} \mid \tau,\psi] = \tfrac{1}{2} \Pr\left\{\varepsilon > \frac{E^2}{\cos^2 \psi \max [m_0^2(-\tau), m_0^2(-\tau - T)]} \,\middle|\, \tau,\psi\right\} \tag{108}$$

and, by (5) this is equal to

$$\Pr[\text{error} \mid \tau,\psi] = \frac{1}{2}\left\{\frac{\varepsilon_0 \cos^2 \psi \max [m_0^2(-\tau), m_0^2(-\tau - T)]}{E^2}\right\}^{\alpha}. \tag{109}$$

Averaging with respect to ψ and τ yields

$$\Pr[\text{error}] = \frac{\beta C_\alpha}{2}\left[\frac{\varepsilon_0}{E^2}\right]^{\alpha} \int_{-\infty}^{\infty} \{\max [m_0^2(-\tau), m_0^2(-\tau - T)]\}^{\alpha} \, d\tau. \tag{110}$$

The signal on the channel $s_c(t)$ is the same as that for the double-sideband AM system. The average transmitted power is therefore given by (21). Substituting this into (110), dividing by βT, and rearranging terms, yields that the conditional error rate is equal to

$$\bar{N} = \frac{C_\alpha}{2}\left[\frac{1}{2} I_1 \frac{\varepsilon_0}{\bar{W}T}\right]^{\alpha} \int_{-\infty}^{\infty} \{\max [Tm_0^2(t), Tm_0^2(t - T)]\}^{\alpha} \frac{dt}{T} \tag{111}$$

where $\bar{W}T/\varepsilon_0$ is the signal-to-noise ratio as defined previously.

In order to obtain numerical results, the general expression for $\bar{N}$ given above is evaluated for the same special case as the preceding systems. The spectrum of a single data pulse is the raised cosine spectrum, and the spectrum of the noise burst is assumed to be constant in the band of interest. The equivalent receiving filter characteristic $H(\omega)$ is the same as that for the AM systems. The expression for $\bar{N}$ has been

numerically evaluated, on a digital computer, for values of α equal to 1, 2, and 3. The results are tabulated below, and are plotted in Fig. 15.

α	$\bar{N}$
1	$0.236\ (\varepsilon_0/\bar{W}T)$
2	$0.109\ (\varepsilon_0/\bar{W}T)^2$
3	$0.0625\ (\varepsilon_0/\bar{W}T)^3$

5.2 *Quaternary (Four Phase) System*

A quaternary phase shift keying system transmits a signal which, at the discrete sampling instants, may have any one of four possible phases spaced 90° apart. Such a signal is mathematically equivalent to two binary signals in quadrature, where each binary signal is of the form described in the preceding section.

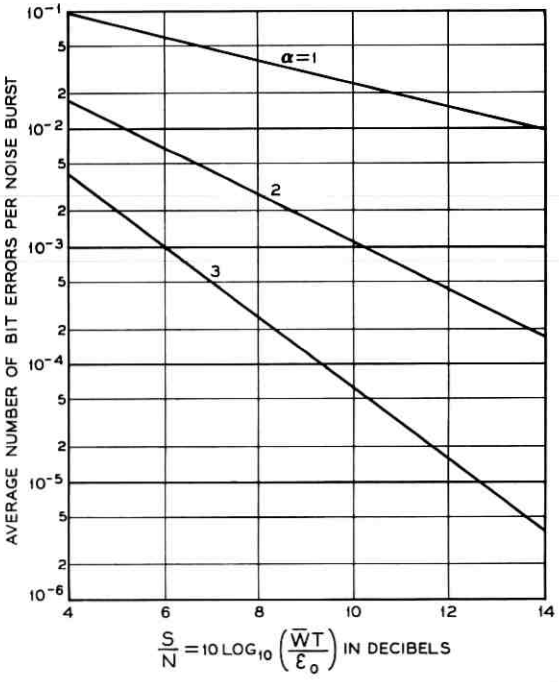


Fig. 15 — Performance of binary phase shift keyed system with differentially coherent detection.

The transmitted signal consists of an "in-phase" binary signal and "quadrature" binary signal. The ideal data receiver consists of a receiving filter $H_c(\omega)$, centered about the carrier frequency ω_c, followed by two separate binary receivers. One of these, identical to the receiver of the preceding section, is sensitive to the "in-phase" binary signal. The other binary receiver, preceded by a phase shifting network with a phase characteristic which is equal to $-\pi/2$ radians throughout the frequency band of the received signal, is sensitive to the "quadrature" binary signal. A block diagram of the receiver is shown in Fig. 16. The system operates in the following manner. In the absence of noise, the signal at the output of the receiving filter is of the form

$$s_r(t) = y_a(t) \cos \omega_c t - y_b(t) \sin \omega_c t, \tag{112}$$

where the modulating signals are each of the form given by (6). At the sampling instants the second term of (112) is equal to zero, and the "a" system is identical to the binary system described in the preceding section. In the absence of noise, the signal at the input to the "b" detector is given by the Hilbert transform

$$\hat{s}(t) = y_b \cos \omega_c t + y_a \sin \omega_c t. \tag{113}$$

At the sampling instants the second term of (113) is equal to zero, and the "b" system is also identical to the binary system. At any arbitrarily chosen sampling instant Pr["a" bit in error] and Pr["b" bit in error] are both given by (110). The average number of bits in error per symbol transmitted is equal to

$$\begin{aligned} \bar{P} = {} & \Pr[\text{"}a\text{" bit in error, "}b\text{" bit correct}] \\ & + \Pr[\text{"}b\text{" bit in error, "}a\text{" bit correct}] \\ & + 2\Pr[\text{"}a\text{" bit and "}b\text{" bit both in error}]. \end{aligned} \tag{114}$$

The Venn diagram in Fig. 17 shows that this is equal to

$$\bar{P} = \Pr[\text{"}a\text{" bit in error}] + \Pr[\text{"}b\text{" bit in error}]. \tag{115}$$

The average number of bits in error per symbol transmitted is therefore,

$$\bar{P} = \beta C_\alpha \left[\frac{\varepsilon_0}{E^2}\right]^\alpha \int_{-\infty}^{\infty} \{\max [m_0^2(-\tau), m_0^2(-\tau - T)]\}^\alpha \, d\tau. \tag{116}$$

Since there are, on the average, βT noise bursts per symbol transmitted, the average number of bits in error per noise burst is equal to

$$\bar{N} = \bar{P}/\beta T. \tag{117}$$

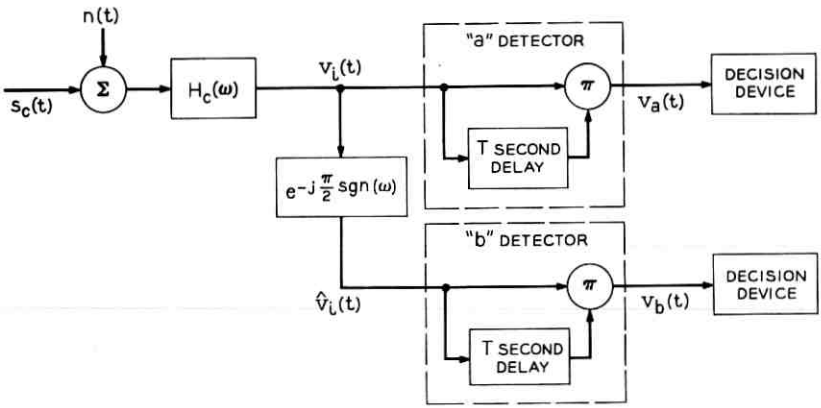


Fig. 16 — Receiver for quaternary phase shift keyed system with differentially coherent detection.

The transmitted signal consists of two binary signals in quadrature. Since the two binary signals are orthogonal, the average power of the transmitted signal is equal to the sum of the average powers of the two binary signals, and is twice the power for the binary system. Substituting this and (117), into (116) and rearranging terms yields

$$\bar{N} = C_\alpha \left[I_1 \frac{\varepsilon_0}{\bar{W}T} \right]^\alpha \int_{-\infty}^{\infty} \{\max [Tm_0^2(t), Tm_0^2(t-T)]\}^\alpha \frac{dt}{T}, \quad (118)$$

where in this case, $\bar{W}T/2\varepsilon_0$ is the signal-to-noise ratio, described previously as the signal energy per bit divided by the minimum energy per noise burst. When $\bar{N}$ is plotted versus 10 $\log_{10}$ $[\bar{W}T/2\varepsilon_0]$ db as the abscissa, the value is twice that for the binary case.

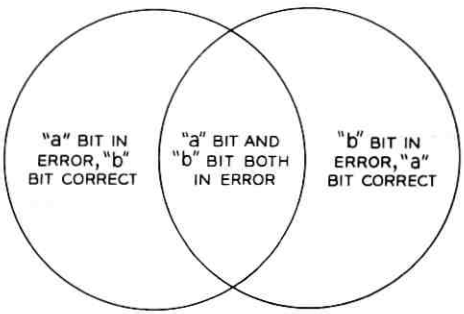


Fig. 17 — Venn diagram for decision outcomes.

VI. COMPARISON OF THE MODULATION SYSTEMS

In the preceding three sections, the various data transmission systems have been analyzed, and the performance of each in the presence of impulsive noise has been determined. It is of interest to compare these results and to rank the systems as to performance. For each of the modulation systems, an expression has been derived for the average number of bit errors per noise burst, as a function of the signal-to-noise ratio. The systems may be ranked by comparing the signal-to-noise ratios required by the different systems for the same error rate.

Such a comparison is done here for the special case for which $\bar{N}$ has been evaluated. The spectrum of an individual data pulse is the raised cosine spectrum, and the spectrum of the noise burst is assumed to be constant in the band of interest. The equivalent receiving filter for each system is the one which minimizes $\bar{N}$ when α is equal to 1. (In those cases for which the optimum filter has impulses in its response, corresponding to carrier suppression at the transmitter, the suboptimum "smooth" filter is assumed. This is described in the preceding sections.)

For each of the modulation systems, the average number of bit errors per noise burst is given by the expression

$$\bar{N} = \left[C \frac{\varepsilon_0}{\bar{W}T} \right]^{\alpha}, \tag{119}$$

where the constant C depends on the type of modulation system, the characteristics of the receiving filter, and the value of α. The systems may be ranked by comparing the values of C. The ratio between the values of C for any two systems is the difference in signal-to-noise ratio required for equal error rate. In Fig. 18, the values of C for the various modulation systems are plotted as functions of α.

The comparison of Fig. 18 shows the modulation systems to be ranked in the following order:

(1.) Single-sideband AM with coherent detection.
(2.) Double-sideband AM with coherent detection.
(3.) Phase shift keying with differentially coherent detection.
(4.) Frequency shift keying.
(5.) AM with envelope detection.

It is interesting to note that this is the same ranking as has been determined for performance in the presence of Gaussian noise.

VII. COMPLEMENTARY DELAY FILTERS

In the sections on the performances of the various modulation systems, the general expressions for the error rate have been evaluated for

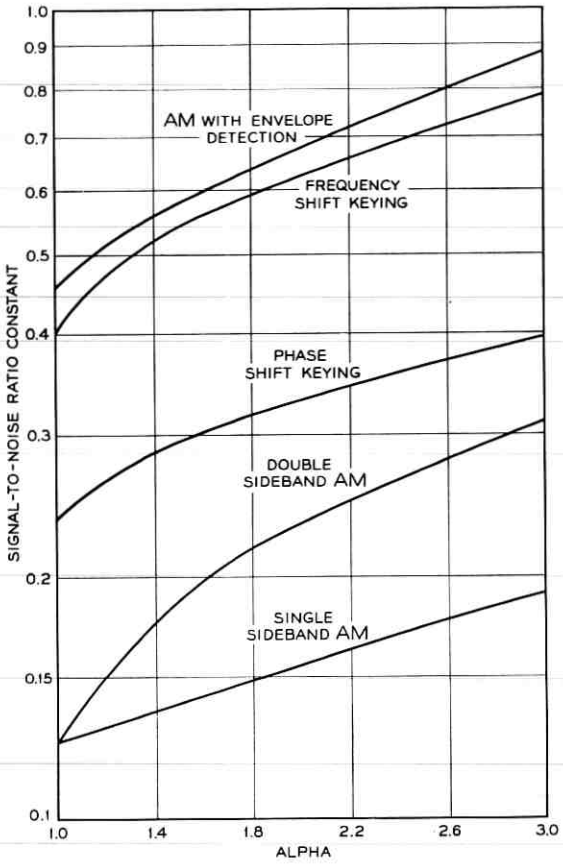


Fig. 18 — Comparison of modulation systems.

a special case which is of wide interest. For the purpose of those evaluations, the equivalent receiving filter $H(\omega)$ has been given an amplitude characteristic $A(\omega)$ which minimizes the error rate for the worst case, when α is equal to 1, and a phase characteristic $\phi(\omega)$ which is zero or a linear function of frequency

$$\phi(\omega) = -D\omega. \tag{120}$$

Such a phase characteristic causes a pure delay of D seconds, with no distortion. Since this corresponds merely to shift of the time axis, and affects signal and noise identically, it has no effect. It has been pointed out previously that, for values of α greater than 1, a phase characteristic

$\phi(\omega)$ which is a function of frequency other than linear can serve to reduce the error rate. This effect is now discussed.

When the phase characteristic is other than a linear function, the envelope delay

$$D(\omega) = -\frac{d}{d\omega}\phi(\omega) \tag{121}$$

is not constant, but varies with frequency. When the noise burst passes through such a filter, its various frequency components are delayed by different amounts, and its energy is spread out over many bits. The peak value of the spread out burst is much lower than the peak value of the original burst, so that many bursts which would have caused errors no longer do so.

In the analyses of the systems, the individual data pulses at the receiving filter output are each constrained to be a specific function $y_0(t)$. This implies that at the transmitter the data pulses pass through a transmitting delay filter, with envelope delay equal to

$$D_T(\omega) = D - D(\omega)$$

in the frequency band of the data signal. The constant D, at least as large as the maximum value of $D(\omega)$, is necessitated by the fact that, for realizability, $D_T(\omega)$ must be positive. The two delay filters exactly complement one another, so that the net effect on the data signal is a pure delay, with no distortion. The noise burst, occurring on the channel after the transmitting filter, only passes through the second filter and is spread out.

The use of such complementary delay filters, based on heuristic reasoning of the type given above, has been suggested previously.[4,11] In this section, the improvement which results from the use of such filters, in terms of the equivalent increase in signal-to-noise ratio which would be required for an equal reduction in error rate, is presented.

Three types of delay networks, which lend themselves to synthesis, have been considered. One of these, called a "linear delay network," has an envelope delay characteristic given by

$$D_1(\omega) = (D_m T/2\pi)\omega \operatorname{sgn}(\omega) \quad \text{for} \quad |\omega| < 2\pi/T \tag{122}$$

and hence a phase characteristic equal to

$$\phi_1(\omega) = -(D_m T/4\pi)\omega^2 \operatorname{sgn}(\omega) \quad \text{for} \quad |\omega| < 2\pi/T, \tag{123}$$

where D_m is the maximum delay in the band of interest. The other two types are called "sinusoidal delay networks." One of these, with a half

cycle of sinusoid in the band of interest has an envelope delay characteristic

$$D_2(\omega) = D_m \cos \omega T/4 \quad \text{for} \quad |\omega| < 2\pi/T \tag{124}$$

and hence a phase characteristic equal to

$$\phi_2(\omega) = -(4D_m/T) \sin \omega T/4 \quad \text{for} \quad |\omega| < 2\pi/T. \tag{125}$$

The third type is a sinusoidal delay network with a full cycle of sinusoid in the band of interest. It has an envelope delay characteristic given by

$$D_3(\omega) = (D_m/2)[1 + \cos \omega T/2] \quad \text{for} \quad |\omega| < 2\pi/T \tag{126}$$

and hence a phase characteristic

$$\phi_3(\omega) = (D_m/T)[(\omega T/2) + \sin \omega T/2] \quad \text{for} \quad |\omega| < 2\pi/T. \tag{127}$$

The three envelope delay chacteristics are shown in Fig. 19.

The general expressions for the error rate, derived previously for the various modulation systems, have been evaluated on a digital computer, with the equivalent receiving filter having each of the above phase characteristics. Values of maximum delay up to 10 symbol durations have been considered. The resulting error rates are compared with those

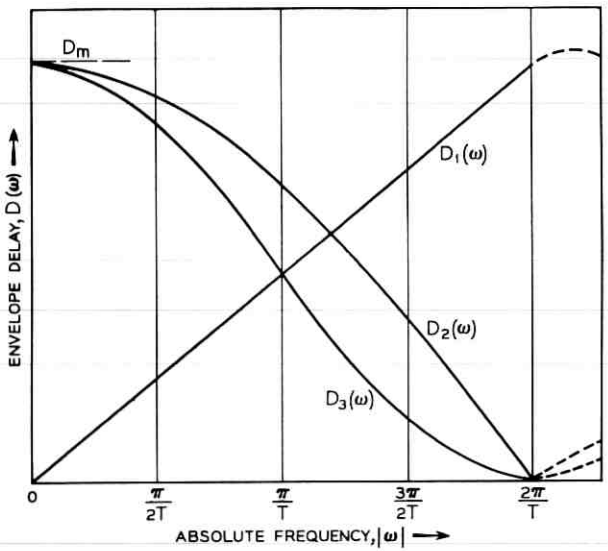


Fig. 19 — Envelope delay characteristics.

when no delay is included, and the improvement is plotted, as a function of the maximum delay, in Figs. 20 through 23.

For all but the differentially coherent phase shift keying systems, the delay networks have no effect for a value of α equal to 1. This has been discussed in the preceding sections; by Parseval's theorem the phase characteristics of the noise have no effect when α is equal to 1. As the value of α increases, the networks become more effective. This phenomenon is to be expected. When the noise burst is spread over many bits, the peak value is reduced and many bursts which would have caused errors no longer do so. However, this improvement is reduced somewhat by the fact that bursts of very large amplitude remain large enough, even after spreading, to cause errors. Such bursts, which would have caused one or two errors, now cause many. As α increases, the percentage of such high amplitude bursts decreases, and the improvement increases. This effect, which reduces the over-all improvement to be gained by the use of complementary delay filters, prompts the consideration of limiting at the input to the delay filter at the receiver. If the limiter is set to a value just above the peak value of the signal, bursts of very high amplitude are clipped and after spreading, will not cause as many errors as they would have without limiting.

A precise analysis of the effect of limiting prior to spreading is not practicable. The process is nonlinear, and the response of the combined limiter-filter depends strongly on the amplitude K and time of occurrence

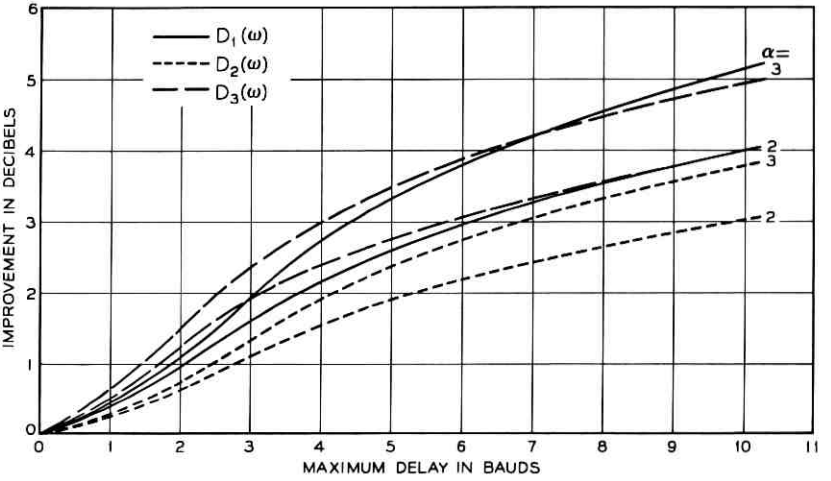


Fig. 20 — Effect of complementary delay networks on double sideband AM systems.

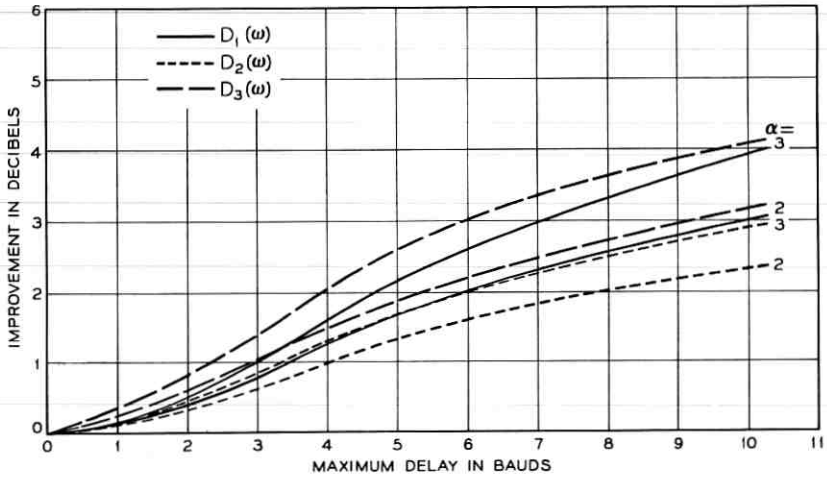


Fig. 21 — Effect of complementary delay networks on single-sideband AM system.

τ of the noise burst, as well as on the particular data sequence being transmitted. In addition, the results of such an analysis would be quite sensitive to the model chosen to represent the channel. In the systems discussed in this paper, the transmission medium and receiving filter have both been linear, and it was therefore possible to combine their

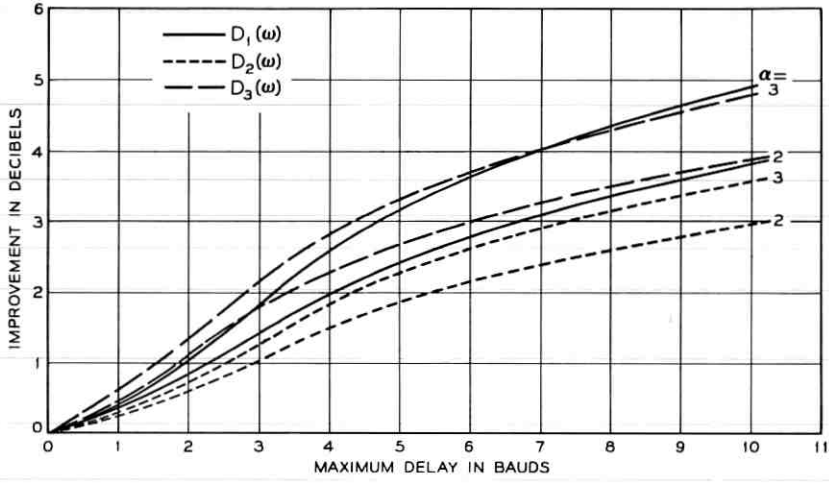


Fig. 22 — Effect of complementary delay networks on frequency shift keyed system.

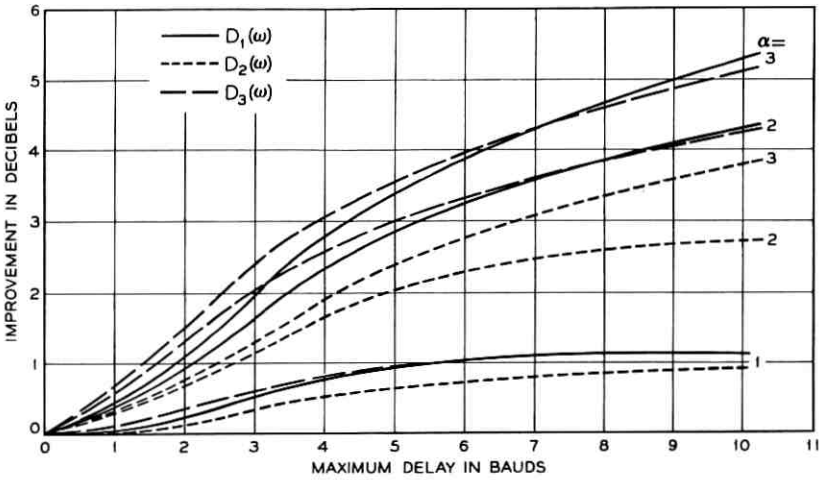


Fig. 23 — Effect of complementary delay networks on differentially coherent phase shift keyed system.

characteristics for the purpose of analysis. The channel was considered to have unity gain in the band of frequencies passed by the receiver band-pass filter; the transmission characteristics of the channel at frequencies outside that band had no effect on the signal or noise at the output of the filter. If limiting is introduced, however, it is most effective if it is done at the point where the noise bursts are most impulsive. This point is at the input to the receiver, prior to any filtering or spreading, where their spectrum is widest. With a nonlinear device between the channel and the receiving filter, their characteristics cannot be combined. The effect of the limiting depends very much on the transmission characteristics of the channel outside the pass band of the filter. If the channel has a considerably wider band than the filter, the limiting will be much more effective than if the bandwidths are comparable.

For these reasons, an analytic evaluation of the improvement resulting from the introduction of limiting is not practicable. An evaluation by simulation techniques or experimental procedures is more feasible. Limiting is discussed in this paper only to point out that further improvement is possible.

VIII. ACKNOWLEDGMENT

This article is taken from the author's Ph.D. dissertation at the Polytechnic Institute of Brooklyn. Thanks are due Prof. M. Schwartz for his assistance as thesis advisor.

The author is particularly indebted to J. Salz, M. A. Rappeport, and A. P. Stamboulis of the Bell Telephone Laboratories. Their comments and assistance have been invaluable.

REFERENCES

1. Alexander, A. A., Gryb, R. M., and Nast, D. W., Capabilities of the Telephone Network for Data Transmission, B.S.T.J., *39*, May, 1960, pp. 431–476.
2. Berger, J. M. and Mandelbrot, B., A New Model for Error Clustering in Telephone Circuits, I.B.M. J. Research and Development, July, 1963.
3. Fennick, J. H., A Report on Some Characteristics of Impulse Noise in Telephone Communication Systems, Conference Paper 63–986, I.E.E.E., June, 1963.
4. Knox-Seith, J., Pulse Transmission, Patent No. 3032725, May 1, 1962.
5. Mertz, P., Model of Impulsive Noise for Data Transmission, I.R.E. Trans. Communications Systems, June, 1961.
6. Mertz, P., Statistics of Hyperbolic Error Distributions in Data Transmission, I.R.E. Trans. Communications Systems, December, 1961.
7. Morris, R., Further Analysis of Errors in Capabilities of the Telephone Network for Data Transmission, B.S.T.J., *41*, July, 1962, pp. 1399–1414.
8. Nyquist, H., Certain Topics in Telegraph Transmission Theory, Trans. A.I.E.E., April, 1928.
9. Sunde, E. D., Ideal Binary Pulse Transmission by AM and FM, B.S.T.J., *38*, Nov., 1959. pp. 1357–1426.
10. Sussman, S. M., Analysis of the Pareto Model for Error Statistics on Telephone Circuits, I.E.E.E. Trans. Communications Systems, June, 1963.
11. Wainwright, R. A., On the Potential Advantage of a Smearing-Desmearing Filter Technique in Overcoming Impulse-Noise Problems in Data Systems, I.R.E. Trans. Communications Systems, December, 1961.

Eigenvalues Associated with Prolate Spheroidal Wave Functions of Zero Order

By DAVID SLEPIAN and ESTELLE SONNENBLICK

(Manuscript received June 10, 1965)

Presented here are tables of values of χ_n *and* λ_n *, quantities defined by the eigenvalue problems*

$$(1 - x^2)\psi_n'' - 2x\psi_n' + (\chi_n - c^2x^2)\psi_n = 0$$

and

$$\lambda_n\psi_n(x) = \int_{-1}^{1} \frac{\sin c(x - y)}{\pi(x - y)} \psi_n(y)dy.$$

In addition, some approximations for these quantities are given and evaluated.

The prolate spheroidal wave functions of zero order, $\psi_n(x)$, $n = 0, 1, \cdots$, are bounded continuous solutions of both the differential equation

$$(1 - x^2)\frac{d^2\psi_n}{dx^2} - 2x\frac{d\psi_n}{dx} + (\chi_n - c^2x^2)\psi_n = 0$$

and the integral equation

$$\lambda_n\psi_n(x) = \int_{-1}^{1} \frac{\sin c(x - y)}{\pi(x - y)} \psi_n(y)dy.$$

The importance of these functions and the corresponding eigenvalues χ_n and λ_n for a great variety of problems, dealing with such diverse matters as lasers, communication theory, optics, noise theory, etc., can be found in the bibliographies of Refs. 1 and 2. It is our purpose here in response to numerous requests to present some numerical values for these eigenvalues.

Tables I and II list values of χ_n and λ_n respectively for $n = 0(1)20(5)40$ and $c = 0(1)20(5)40$. The values given are, we believe, accurate to all eight figures listed.* Results of the computation are shown graphically on Figs. 1–4.

The values of χ_n were obtained using the method of Bouwkamp as explained for example in Flammer.[1] This computation also gives expansion coefficients $d_r^{on}(c)$ in Flammer's notation, from which his quantity $R_{on}^{(1)}(c,1)$ can be computed. The λ's where then found from

$$\lambda_n = \frac{2c}{\pi}[R_{on}^{(1)}(c,1)]^2.$$

The tables presented required 0.027 hours of computing time on the IBM 7090.

The following formulae for λ_n and χ_n are given in Ref. 2. For fixed n and small c

$$\lambda_n = \frac{2}{\pi}\left[\frac{2^{2n}(n!)^3}{(2n)!(2n+1)!}\right]^2 c^{2n+1} \cdot\left[1 - \frac{(2n+1)c^2}{(2n-1)^2(2n+3)^2} + 0(c^4)\right]. \tag{1}$$

For fixed n and large c

$$1 - \lambda_n = \frac{2^{3n+2}\sqrt{\pi}c^{n+\frac{1}{2}}e^{-2c}}{n!}\left[1 - \frac{6n^2 - 2n + 3}{32c} + 0\left(\frac{1}{c^2}\right)\right]. \tag{2}$$

Some values computed from the terms explicitly exhibited in (1) and (2) are shown as dotted lines on Figs. 3 and 4.

For n and c both large, we have the following result. Let b be fixed and let

$$n = \left[\frac{2}{\pi}(c + b \ln 2\sqrt{c})\right] \tag{3}$$

where the brackets denote "integer part of". Then

$$\lim_{c\to\infty} \lambda_n = (1 + e^{\pi b})^{-1}. \tag{4}$$

* The notation $E \pm XY$ following an entry in the tables indicates that the entry is to be multiplied by $10^{\pm XY}$ where XY is an integer in decimal notation, e.g., $E + 03$ denotes a factor of 10^3.

TABLE I—χ_n

n \ c =	1.	2.	3.	4.	5.	6.
0	3.1900006E − 01	1.1277341E + 00	2.1367322E + 00	3.1720674E + 00	4.1951289E + 00	5.2082692E + 00
1	2.5930846E + 00	4.2871285E + 00	6.8208883E + 00	9.8059438E + 00	1.2911703E + 01	1.6000443E + 01
2	6.5334718E + 00	8.2257130E + 00	1.1192939E + 01	1.5306300E + 01	2.0176915E + 01	2.5356479E + 01
3	1.2514462E + 01	1.4100204E + 01	1.6889030E + 01	2.1048961E + 01	2.6587360E + 01	3.3204199E + 01
4	2.0508274E + 01	2.2054830E + 01	2.4708535E + 01	2.8596855E + 01	3.3897096E + 01	4.0720194E + 01
5	3.0505405E + 01	3.2035263E + 01	3.4631281E + 01	3.8367138E + 01	4.3358996E + 01	4.9773712E + 01
6	4.2503818E + 01	4.4024748E + 01	4.6591428E + 01	5.0252698E + 01	5.5080962E + 01	6.1180757E + 01
7	5.6502845E + 01	5.8018371E + 01	6.0567636E + 01	6.4186116E + 01	6.8924773E + 01	7.4852867E + 01
8	7.2502203E + 01	7.4014194E + 01	7.6552160E + 01	8.0143235E + 01	8.4825931E + 01	9.0651159E + 01
9	9.0501757E + 01	9.2011304E + 01	9.4541490E + 01	9.8113806E + 01	1.0275858E + 02	1.0851545E + 02
10	1.1050143E + 02	1.1200922E + 02	1.1453381E + 02	1.1809267E + 02	1.2271039E + 02	1.2841888E + 02
11	1.3250119E + 02	1.3400766E + 02	1.3652809E + 02	1.4007696E + 02	1.4467463E + 02	1.5034744E + 02
12	1.5650101E + 02	1.5800647E + 02	1.6052372E + 02	1.6406496E + 02	1.6864733E + 02	1.7429300E + 02
13	1.8250086E + 02	1.8400554E + 02	1.8652029E + 02	1.9005557E + 02	1.9462601E + 02	2.0025051E + 02
14	2.1050075E + 02	2.1200480E + 02	2.1451756E + 02	2.1804808E + 02	2.2260902E + 02	2.2821669E + 02
15	2.4050065E + 02	2.4200419E + 02	2.4451535E + 02	2.4804202E + 02	2.5259526E + 02	2.5818931E + 02
16	2.7250058E + 02	2.7400370E + 02	2.7651353E + 02	2.8003704E + 02	2.8458396E + 02	2.9016684E + 02
17	3.0650051E + 02	3.0800328E + 02	3.1051202E + 02	3.1403289E + 02	3.1857456E + 02	3.2414815E + 02
18	3.4250046E + 02	3.4400294E + 02	3.4651075E + 02	3.5002941E + 02	3.5456666E + 02	3.6013245E + 02
19	3.8050041E + 02	3.8200264E + 02	3.8450967E + 02	3.8802645E + 02	3.9255996E + 02	3.9811912E + 02
20	4.2050037E + 02	4.2200239E + 02	4.2450874E + 02	4.2802392E + 02	4.3255422E + 02	4.3810771E + 02
25	6.5050024E + 02	6.5200154E + 02	6.5450564E + 02	6.5801543E + 02	6.6253497E + 02	6.6806946E + 02
30	9.3050017E + 02	9.3200108E + 02	9.3450394E + 02	9.3801078E + 02	9.4252442E + 02	9.4804850E + 02
35	1.2605001E + 03	1.2620008E + 03	1.2645029E + 03	1.2680079E + 03	1.2725180E + 03	1.2780358E + 03
40	1.6405001E + 03	1.6420006E + 03	1.6445022E + 03	1.6480061E + 03	1.6525138E + 03	1.6580275E + 03

TABLE I — *Continued*

n \ c =	7.	8.	9.	10.	11.	12.
0	6.2162529E + 00	7.2215789E + 00	8.2254064E + 00	9.2283043E + 00	1.0230581E + 01	1.1232421E + 01
1	1.9056678E + 01	2.2092154E + 01	2.5116120E + 01	2.8133464E + 01	3.1146682E + 01	3.4157135E + 01
2	3.0560201E + 01	3.5706417E + 01	4.0802950E + 01	4.5868953E + 01	5.0916879E + 01	5.5953514E + 01
3	4.0405727E + 01	4.7757099E + 01	5.5051178E + 01	6.2257700E + 01	6.9401323E + 01	7.6505824E + 01
4	4.8910585E + 01	5.8016770E + 01	6.7500818E + 01	7.6993289E + 01	8.6369707E + 01	9.5638659E + 01
5	5.7777751E + 01	6.7364750E + 01	7.8205025E + 01	8.9739267E + 01	1.0144734E + 02	1.1305411E + 02
6	6.8701439E + 01	7.7825223E + 01	8.8658000E + 01	1.0103543E + 02	1.1446976E + 02	1.2835139E + 02
7	8.2064637E + 01	9.0691430E + 01	1.0090790E + 02	1.1288107E + 02	1.2660565E + 02	1.4174147E + 02
8	9.7684571E + 01	1.0601169E + 02	1.1574796E + 02	1.2705083E + 02	1.4010628E + 02	1.5502454E + 02
9	1.1543428E + 02	1.2357716E + 02	1.3302232E + 02	1.4387201E + 02	1.5626473E + 02	1.7038033E + 02
10	1.3525788E + 02	1.4327579E + 02	1.5253134E + 02	1.6309665E + 02	1.7506323E + 02	1.8855245E + 02
11	1.5712799E + 02	1.6505554E + 02	1.7417688E + 02	1.8454762E + 02	1.9623470E + 02	2.0932095E + 02
12	1.8102925E + 02	1.8888879E + 02	1.9791014E + 02	2.0813839E + 02	2.1962634E + 02	2.3243647E + 02
13	2.0695232E + 02	2.1475915E + 02	2.2370349E + 02	2.3382295E + 02	2.4516097E + 02	2.5776775E + 02
14	2.3489114E + 02	2.4265622E + 02	2.5153975E + 02	2.6157378E + 02	2.7279496E + 02	2.8524510E + 02
15	2.6484166E + 02	2.7257305E + 02	2.8140764E + 02	2.9137313E + 02	3.0250101E + 02	3.1482692E + 02
16	2.9680105E + 02	3.0450484E + 02	3.1329940E + 02	3.2320895E + 02	3.3426093E + 02	3.4648622E + 02
17	3.3076730E + 02	3.3844819E + 02	3.4720956E + 02	3.5707281E + 02	3.6806210E + 02	3.8020452E + 02
18	3.6673895E + 02	3.7440061E + 02	3.8313413E + 02	3.9295859E + 02	4.0389544E + 02	4.1596869E + 02
19	4.0471489E + 02	4.1236024E + 02	4.2107018E + 02	4.3086179E + 02	4.4175430E + 02	4.5376914E + 02
20	4.4469430E + 02	4.5232571E + 02	4.6101548E + 02	4.7077902E + 02	4.8163367E + 02	4.9359870E + 02
25	6.7462529E + 02	6.8220999E + 02	6.9083226E + 02	7.0050200E + 02	7.1123026E + 02	7.2302932E + 02
30	9.5458747E + 02	9.6214660E + 02	9.7073194E + 02	9.8035039E + 02	9.9100965E + 02	1.0027182E + 03
35	1.2845645E + 03	1.2921081E + 03	1.3006711E + 03	1.3102584E + 03	1.3208759E + 03	1.3325297E + 03
40	1.6645496E + 03	1.6720830E + 03	1.6806314E + 03	1.6901985E + 03	1.7007886E + 03	1.7124067E + 03

TABLE I — *Continued*

n \ c =	13.	14.	15.	16.	17.	18.
0	1.2233939E + 01	1.3235214E + 01	1.4236300E + 01	1.5237237E + 01	1.6238054E + 01	1.7238772E + 01
1	3.7165631E + 01	4.0172681E + 01	4.3178630E + 01	4.6183721E + 01	4.9188129E + 01	5.2191983E + 01
2	6.0982596E + 01	6.6006328E + 01	7.1026104E + 01	7.6042858E + 01	8.1057244E + 01	8.6069739E + 01
3	8.3585719E + 01	9.0649230E + 01	9.7701181E + 01	1.0474459E + 02	1.1178148E + 02	1.1881324E + 02
4	1.0483494E + 02	1.1398465E + 02	1.2310348E + 02	1.3220071E + 02	1.4128205E + 02	1.5035127E + 02
5	1.2450713E + 02	1.3584050E + 02	1.4709398E + 02	1.5829441E + 02	1.6945806E + 02	1.8059490E + 02
6	1.4223009E + 02	1.5593166E + 02	1.6945804E + 02	1.8285785E + 02	1.9617229E + 02	2.0942797E + 02
7	1.5768187E + 02	1.7382917E + 02	1.8983663E + 02	2.0562050E + 02	2.2121898E + 02	2.3668721E + 02
8	1.7164637E + 02	1.8946449E + 02	2.0780930E + 02	2.2615098E + 02	2.4425790E + 02	2.6211708E + 02
9	1.8639078E + 02	2.0430123E + 02	2.2377440E + 02	2.4417004E + 02	2.6481922E + 02	2.8529987E + 02
10	2.0372711E + 02	2.2077813E + 02	2.3982549E + 02	2.6072449E + 02	2.8296138E + 02	3.0581686E + 02
11	2.2391335E + 02	2.4015533E + 02	2.5823139E + 02	2.7832097E + 02	3.0044546E + 02	3.2428066E + 02
12	2.4664450E + 02	2.6234574E + 02	2.7966575E + 02	2.9877161E + 02	3.1985976E + 02	3.4306306E + 02
13	2.7170195E + 02	2.8703335E + 02	3.0384785E + 02	3.2225590E + 02	3.4240470E + 02	3.6448308E + 02
14	2.9897207E + 02	3.1403113E + 02	3.3048712E + 02	3.4841822E + 02	3.6792260E + 02	3.8912905E + 02
15	3.2839115E + 02	3.4323943E + 02	3.5942402E + 02	3.7700550E + 02	3.9605573E + 02	4.1666293E + 02
16	3.5991943E + 02	3.7459944E + 02	3.9056997E + 02	4.0788062E + 02	4.2658830E + 02	4.4675957E + 02
17	3.9353027E + 02	4.0807299E + 02	4.2387018E + 02	4.4096376E + 02	4.5940093E + 02	4.7923537E + 02
18	4.2920498E + 02	4.4363385E + 02	4.5928798E + 02	4.7620357E + 02	4.9442089E + 02	5.1398496E + 02
19	4.6693004E + 02	4.8126322E + 02	4.9679753E + 02	5.1356471E + 02	5.3159977E + 02	5.5094143E + 02
20	5.0669542E + 02	5.2094728E + 02	5.3638001E + 02	5.5302178E + 02	5.7090346E + 02	5.9005891E + 02
25	7.3591267E + 02	7.4989504E + 02	7.6499245E + 02	7.8122221E + 02	7.9860305E + 02	8.1715510E + 02
30	1.0154855E + 03	1.0293217E + 03	1.0442378E + 03	1.0602457E + 03	1.0773582E + 03	1.0955890E + 03
35	1.3452267E + 03	1.3589745E + 03	1.3737811E + 03	1.3896552E + 03	1.4066063E + 03	1.4246442E + 03
40	1.7250580E + 03	1.7387482E + 03	1.7534835E + 03	1.7692707E + 03	1.7861168E + 03	1.8040295E + 03

TABLE I — *Continued*

n \ c =	19.	20.	25.	30.	35.	40.
0	1.8239408E + 01	1.9239976E + 01	2.4242094E + 01	2.9243472E + 01	3.4244440E + 01	3.9245159E + 01
1	5.5195383E + 01	5.8198404E + 01	7.3209570E + 01	8.8216755E + 01	1.0322177E + 02	1.1822547E + 02
2	9.1080697E + 01	9.6090388E + 01	1.2112584E + 02	1.4614836E + 02	1.7116395E + 02	1.9617538E + 02
3	1.2584090E + 02	1.3286522E + 02	1.6795309E + 02	2.0300813E + 02	2.3804589E + 02	2.7307342E + 02
4	1.5941100E + 02	1.6846310E + 02	2.1364862E + 02	2.5876280E + 02	3.0384036E + 02	3.4889654E + 02
5	1.9171143E + 02	2.0281205E + 02	2.5816358E + 02	3.1337546E + 02	3.6851770E + 02	4.2361994E + 02
6	2.2264133E + 02	2.3582286E + 02	3.0144041E + 02	3.6680477E + 02	4.3204536E + 02	4.9721681E + 02
7	2.5206574E + 02	2.6738042E + 02	3.4341554E + 02	4.1900403E + 02	4.9438749E + 02	5.6965810E + 02
8	2.7978768E + 02	2.9732623E + 02	3.8400598E + 02	4.6991994E + 02	5.5550429E + 02	6.4091212E + 02
9	3.0548786E + 02	3.2541914E + 02	4.2311389E + 02	5.1949088E + 02	6.1535127E + 02	7.1094418E + 02
10	3.2868309E + 02	3.5126388E + 02	4.6061231E + 02	5.6764441E + 02	6.7387820E + 02	7.7971605E + 02
11	3.4916478E + 02	3.7436419E + 02	4.9632769E + 02	6.1429371E + 02	7.3102784E + 02	8.4718535E + 02
12	3.6825795E + 02	3.9493519E + 02	5.2999995E + 02	6.5933191E + 02	7.8673407E + 02	9.1330472E + 02
13	3.8867917E + 02	4.1503422E + 02	5.6120901E + 02	7.0262235E + 02	8.4091938E + 02	9.7802077E + 02
14	4.1220822E + 02	4.3736223E + 02	5.8939527E + 02	7.4397983E + 02	8.9349109E + 02	1.0412727E + 03
15	4.3894029E + 02	4.6303801E + 02	6.1457579E + 02	7.8313074E + 02	9.4433551E + 02	1.1029905E + 03
16	4.6847456E + 02	4.9183371E + 02	6.3866220E + 02	8.1963882E + 02	9.9330804E + 02	1.1630922E + 03
17	5.0052914E + 02	5.2335572E + 02	6.6469676E + 02	8.5289508E + 02	1.0402144E + 03	1.2214800E + 03
18	5.3494661E + 02	5.5736402E + 02	6.9431366E + 02	8.8272216E + 02	1.0847715E + 03	1.2780344E + 03
19	5.7163273E + 02	5.9372195E + 02	7.2743245E + 02	9.1079959E + 02	1.1265337E + 03	1.3326043E + 03
20	6.1052541E + 02	6.3234422E + 02	7.6357631E + 02	9.4039863E + 02	1.1648560E + 03	1.3849884E + 03
25	8.3690003E + 02	8.5786114E + 02	9.8185790E + 02	1.1410460E + 03	1.3423987E + 03	1.6007929E + 03
30	1.1149527E + 03	1.1354649E + 03	1.2558699E + 03	1.4079656E + 03	1.5952330E + 03	1.8229611E + 03
35	1.4437796E + 03	1.4640237E + 03	1.5823180E + 03	1.7304322E + 03	1.9106676E + 03	2.1260735E + 03
40	1.8230168E + 03	1.8430874E + 03	1.9600239E + 03	2.1056213E + 03	2.2815621E + 03	2.4899642E + 03

TABLE II — λ_n

n \ c =	1.	2.	3.	4.	5.	6.
0	5.7258178E − 01	8.8055992E − 01	9.7582863E − 01	9.9588549E − 01	9.9935241E − 01	9.9990188E − 01
1	6.2791274E − 02	3.5564063E − 01	7.0996324E − 01	9.1210742E − 01	9.7986456E − 01	9.9606164E − 01
2	1.2374793E − 03	3.5867688E − 02	2.0513868E − 01	5.1905484E − 01	7.9992193E − 01	9.4017339E − 01
3	9.2009770E − 06	1.1522328E − 03	1.8203800E − 02	1.1021099E − 01	3.4356219E − 01	6.4679195E − 01
4	3.7179286E − 08	1.8881549E − 05	7.0814710E − 04	8.8278764E − 03	5.6015851E − 02	2.0734922E − 01
5	9.4914367E − 11	1.9358522E − 07	1.6551244E − 05	3.8129172E − 04	4.1820948E − 03	2.7387166E − 02
6	1.6715716E − 13	1.3660608E − 09	2.6410165E − 07	1.0950871E − 05	1.9330846E − 04	1.9550007E − 03
7	2.1544491E − 16	7.0488855E − 12	3.0737365E − 09	2.2786389E − 07	6.3591502E − 06	9.4848766E − 05
8	2.1207239E − 19	2.7767898E − 14	2.7281307E − 11	3.6065493E − 09	1.5822998E − 07	3.4367833E − 06
9	1.6466214E − 22	8.6266788E − 17	1.9085689E − 13	4.4938297E − 11	3.0917257E − 09	9.7321160E − 08
10	1.0343492E − 25	2.1680119E − 19	1.0797906E − 15	4.5252285E − 13	4.8757393E − 11	2.2189805E − 09
11	5.3650197E − 29	4.4986573E − 22	5.0431156E − 18	3.7603029E − 15	6.3402794E − 13	4.1662263E − 11
12	2.3367231E − 32	7.8382450E − 25	1.9775436E − 20	2.6228187E − 17	6.9173022E − 15	6.5574786E − 13
13	8.6674831E − 36	1.1630367E − 27	6.6033063E − 23	1.5575942E − 19	6.4235507E − 17	8.7803771E − 15
14	2.7709612E − 39	1.4873466E − 30	1.9002929E − 25	7.9711081E − 22	5.1393068E − 19	1.0125783E − 16
15	0.	1.6563614E − 33	4.7620029E − 28	3.5519080E − 24	3.5797463E − 21	1.0163838E − 18
16	0.	1.6207613E − 36	1.0485031E − 30	1.3905716E − 26	2.1905130E − 23	8.9610464E − 21
17	0.	0.	2.0444867E − 33	4.8210691E − 29	1.1869344E − 25	6.9950907E − 23
18	0.	0.	3.5551880E − 36	1.4905449E − 31	5.7350388E − 28	4.8687451E − 25
19	0.	0.	5.5475853E − 39	4.1352414E − 34	2.4864675E − 30	3.0405184E − 27
20	0.	0.	0.	1.0352225E − 36	9.7273155E − 33	1.7132439E − 29
25	0.	0.	0.	0.	0.	0.
30	0.	0.	0.	0.	0.	0.
35	0.	0.	0.	0.	0.	0.
40	0.	0.	0.	0.	0.	0.

TABLE II — *Continued*

n \ c =	7.	8.	9.	10.	11.	12.
0	9.9998546E − 01	9.9999787E − 01	9.9999969E − 01	9.9999996E − 01	9.9999999E − 01	1.0000000E + 00
1	9.9929217E − 01	9.9987898E − 01	9.9997999E − 01	9.9999677E − 01	9.9999949E − 01	9.9999992E − 01
2	9.8570806E − 01	9.9700462E − 01	9.9941873E − 01	9.9989273E − 01	9.9998091E − 01	9.9999670E − 01
3	8.6456615E − 01	9.6054568E − 01	9.9039622E − 01	9.9790124E − 01	9.9957158E − 01	9.9991663E − 01
4	4.7705272E − 01	7.4790284E − 01	9.1013316E − 01	9.7445778E − 01	9.9371700E − 01	9.9858732E − 01
5	1.1572386E − 01	3.2027663E − 01	5.9909617E − 01	8.2514635E − 01	9.4136927E − 01	9.8366430E − 01
6	1.3055972E − 02	6.0784427E − 02	1.9693935E − 01	4.4015011E − 01	7.0394130E − 01	8.8175663E − 01
7	9.0657300E − 04	6.1262894E − 03	3.0565075E − 02	1.1232482E − 01	2.9607849E − 01	5.5736081E − 01
8	4.5623948E − 05	4.1825206E − 04	2.8466070E − 03	1.4920175E − 02	6.0370339E − 02	1.8342927E − 01
9	1.7774751E − 06	2.1663088E − 05	1.9230822E − 04	1.3145890E − 03	7.1417030E − 03	3.1054179E − 02
10	5.5526131E − 08	8.9304272E − 07	1.0194316E − 05	8.8213430E − 05	6.0469421E − 04	3.3745471E − 03
11	1.4251398E − 09	3.0137350E − 08	4.3973999E − 07	4.7664454E − 06	4.0395675E − 05	2.7741888E − 04
12	3.0622379E − 11	8.4965846E − 10	1.5795600E − 08	2.1339628E − 07	2.2179166E − 06	1.8475085E − 05
13	5.5928434E − 13	2.0334083E − 11	4.8068821E − 10	8.0707164E − 09	1.0243298E − 07	1.0282524E − 06
14	8.7926605E − 15	4.1852675E − 13	1.2564804E − 11	2.6170188E − 10	4.0455355E − 09	4.8758791E − 08
15	1.2026890E − 16	7.4905020E − 15	2.8533973E − 13	7.3634903E − 12	1.3840557E − 10	1.9981456E − 09
16	1.4445726E − 18	1.1767148E − 16	5.6843266E − 15	1.8159383E − 13	4.1453619E − 12	7.1571886E − 11
17	1.5359357E − 20	1.6358709E − 18	1.0016699E − 16	3.9589753E − 15	1.0966649E − 13	2.2619074E − 12
18	1.4559023E − 22	2.0270123E − 20	1.5727550E − 18	7.6870812E − 17	2.5823710E − 15	6.3575326E − 14
19	1.2380854E − 24	2.2529462E − 22	2.2145250E − 20	1.3380681E − 18	5.4488496E − 17	1.6002320E − 15
20	9.4989023E − 27	2.2588880E − 24	2.8123556E − 22	2.1001719E − 20	1.0363386E − 18	3.6290304E − 17
25	6.3410693E − 38	5.7412040E − 35	2.3265268E − 32	4.9987005E − 30	6.4253487E − 28	5.4015219E − 26
30	0.	0.	0.	0.	5.4863023E − 38	1.1037482E − 35
35	0.	0.	0.	0.	0.	0.
40	0.	0.	0.	0.	0.	0.

TABLE II — *Continued*

n \ c =	13.	14.	15.	16.	17.	18.
0	1.0000000E + 00	1.0000000E + 00	1.0000000E + 00	1.0000000E + 00	1.0000000E + 00	1.0000000E + 00
1	9.9999999E − 01	1.0000000E + 00	1.0000000E + 00	1.0000000E + 00	1.0000000E + 00	1.0000000E + 00
2	9.9999944E − 01	9.9999991E − 01	9.9999998E − 01	1.0000000E + 00	1.0000000E + 00	1.0000000E + 00
3	9.9998436E − 01	9.9999715E − 01	9.9999949E − 01	9.9999991E − 01	9.9999998E − 01	1.0000000E + 00
4	9.9970100E − 01	9.9993948E − 01	9.9998818E − 01	9.9999776E − 01	9.9999958E − 01	9.9999992E − 01
5	9.9595266E − 01	9.9907074E − 01	9.9979787E − 01	9.9995783E − 01	9.9999149E − 01	9.9999833E − 01
6	9.6225505E − 01	9.8963945E − 01	9.9741834E − 01	9.9939756E − 01	9.9986611E − 01	9.9997138E − 01
7	7.8874478E − 01	9.2170099E − 01	9.7594492E − 01	9.9346756E − 01	9.9836416E − 01	9.9961326E − 01
8	4.0633176E − 01	6.6365081E − 01	8.5371077E − 01	9.4900699E − 01	9.8478345E − 01	9.9589845E − 01
9	1.0588399E − 01	2.7254759E − 01	5.1899118E − 01	7.5367260E − 01	9.0120813E − 01	9.6721066E − 01
10	1.5487268E − 02	5.7772021E − 02	1.6922485E − 01	3.7484512E − 01	6.2548476E − 01	8.2543081E − 01
11	1.5807741E − 03	7.5604029E − 03	3.0214721E − 02	9.8343344E − 02	2.5013211E − 01	4.8298471E − 01
12	1.2703997E − 04	7.3609236E − 04	3.6365712E − 03	1.5325905E − 02	5.4203716E − 02	1.5521339E − 01
13	8.4414630E − 06	5.8099164E − 05	3.4130591E − 04	1.7310585E − 03	7.6061834E − 03	2.8698205E − 02
14	4.7534082E − 07	3.8540665E − 06	2.6544166E − 05	1.5775571E − 04	8.1779798E − 04	3.7164510E − 03
15	2.3057257E − 08	2.1923715E − 07	1.7585578E − 06	1.2118149E − 05	7.2740820E − 05	3.8416435E − 04
16	9.7549162E − 10	1.0846150E − 08	1.0092888E − 07	8.0200884E − 07	5.5289258E − 06	3.3476998E − 05
17	3.6359074E − 11	4.7180556E − 10	5.0802489E − 09	4.6393027E − 08	3.6561510E − 07	2.5213492E − 06
18	1.2039857E − 12	1.8208273E − 11	2.2646216E − 10	2.3711707E − 09	2.1297926E − 08	1.6661751E − 07
19	3.5675475E − 14	6.2816688E − 13	9.0124935E − 12	1.0801803E − 10	1.1034773E − 09	9.7669046E − 09
20	9.5186419E − 16	1.9498859E − 14	3.2240722E − 13	4.4178809E − 12	5.1252279E − 11	5.1225654E − 10
25	3.1773164E − 24	1.3781710E − 22	4.5947254E − 21	1.2173462E − 19	2.6337453E − 18	4.7585812E − 17
30	1.4501905E − 33	1.3251429E − 31	8.8515118E − 30	4.4989186E − 28	1.7978779E − 26	5.8040997E − 25
35	0.	0.	0.	3.1590168E − 37	2.3233738E − 35	1.3343537E − 33
40	0.	0.	0.	0.	0.	0.

TABLE II — *Continued*

n \ c =	19.	20.	25.	30.	35.	40.
0	1.0000000E + 00	1.0000000E + 00	1.0000000E + 00	1.0000000E + 00	1.0000000E + 00	1.0000000E + 00
1	1.0000000E + 00	1.0000000E + 00	1.0000000E + 00	1.0000000E + 00	1.0000000E + 00	1.0000000E + 00
2	1.0000000E + 00	1.0000000E + 00	1.0000000E + 00	1.0000000E + 00	1.0000000E + 00	1.0000000E + 00
3	1.0000000E + 00	1.0000000E + 00	1.0000000E + 00	1.0000000E + 00	1.0000000E + 00	1.0000000E + 00
4	9.9999999E − 01	1.0000000E + 00	1.0000000E + 00	1.0000000E + 00	1.0000000E + 00	1.0000000E + 00
5	9.9999968E − 01	9.9999994E − 01	1.0000000E + 00	1.0000000E + 00	1.0000000E + 00	1.0000000E + 00
6	9.9999408E − 01	9.9999881E − 01	1.0000000E + 00	1.0000000E + 00	1.0000000E + 00	1.0000000E + 00
7	9.9991254E − 01	9.9998093E − 01	9.9999999E − 01	1.0000000E + 00	1.0000000E + 00	1.0000000E + 00
8	9.9896831E − 01	9.9975345E − 01	9.9999988E − 01	1.0000000E + 00	1.0000000E + 00	1.0000000E + 00
9	9.9042654E − 01	9.9743251E − 01	9.9999821E − 01	1.0000000E + 00	1.0000000E + 00	1.0000000E + 00
10	9.3461880E − 01	9.7911569E − 01	9.9997682E − 01	9.9999999E − 01	1.0000000E + 00	1.0000000E + 00
11	7.1923718E − 01	8.7971361E − 01	9.9974565E − 01	9.9999983E − 01	1.0000000E + 00	1.0000000E + 00
12	3.4534703E − 01	5.8879338E − 01	9.9766185E − 01	9.9999783E − 01	1.0000000E + 00	1.0000000E + 00
13	9.0528307E − 02	2.2898871E − 01	9.8251216E − 01	9.9997547E − 01	9.9999998E − 01	1.0000000E + 00
14	1.4751648E − 02	5.0245996E − 02	9.0214476E − 01	9.9975907E − 01	9.9999980E − 01	1.0000000E + 00
15	1.7952937E − 03	7.4212338E − 03	6.5129574E − 01	9.9796698E − 01	9.9999766E − 01	1.0000000E + 00
16	1.7967267E − 04	8.5983868E − 04	2.9167771E − 01	9.8564508E − 01	9.9997604E − 01	9.9999998E − 01
17	1.5383334E − 05	8.3739541E − 05	7.5468799E − 02	9.2101083E − 01	9.9978298E − 01	9.9999978E − 01
18	1.1493460E − 06	7.0600702E − 06	1.3031043E − 02	7.0692287E − 01	9.9828070E − 01	9.9999766E − 01
19	7.5908731E − 08	5.2374892E − 07	1.7588754E − 03	3.5647890E − 01	9.8836235E − 01	9.9997777E − 01
20	4.4748828E − 09	3.4574493E − 08	2.0082884E − 04	1.0627740E − 01	9.3662832E − 01	9.9981076E − 01
25	7.3154280E − 16	9.7203030E − 15	6.7594641E − 10	4.7379571E − 06	5.3273087E − 03	4.8731168E − 01
30	1.5483211E − 23	3.4788064E − 22	2.4882077E − 16	1.3177338E − 11	1.0489720E − 07	1.7876070E − 04
35	6.1447670E − 32	2.3202540E − 30	1.6231560E − 23	5.8407336E − 18	2.5646331E − 13	2.2391506E − 09
40	0.	0.	2.4836333E − 31	5.8205690E − 25	1.3022033E − 19	4.9862021E − 15

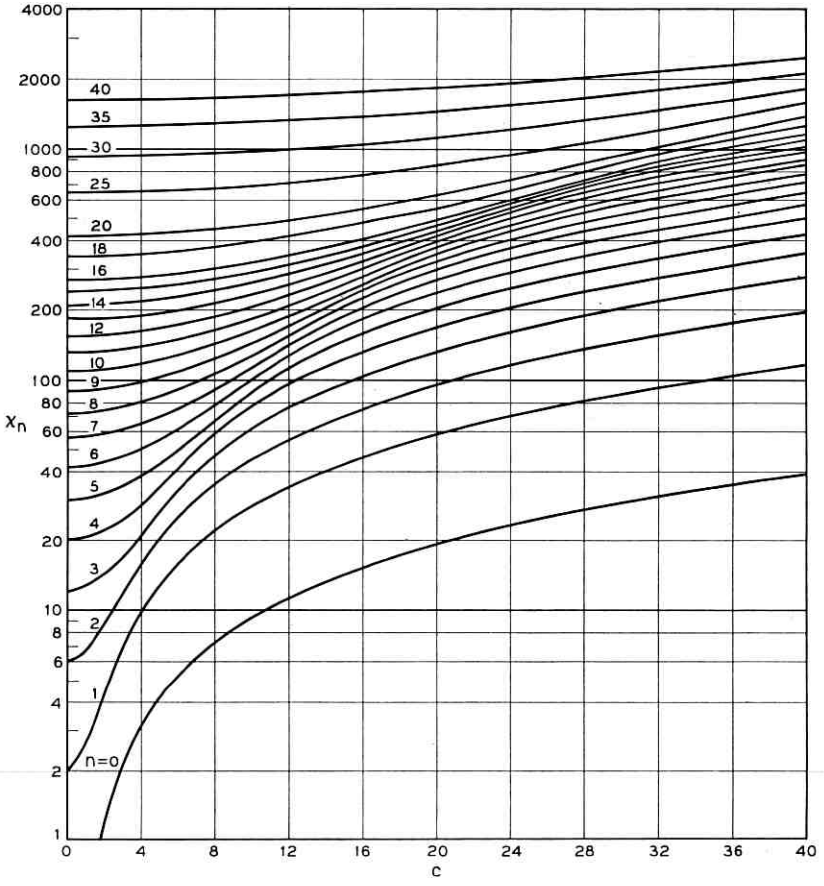


Fig. 1 — Eigenvalues, χ_n, of $(1 - x^2)\psi'' - 2x\psi' + (\chi - c^2x^2)\psi = 0$.

The derivation of (3) and (4) given in Ref. 2 suggests the approximate formula

$$\lambda_n \approx \hat{\lambda}_n = (1 + e^{\pi\hat{b}})^{-1} \tag{5}$$

$$\hat{b} = \frac{n\frac{\pi}{2} - c + \frac{\pi}{4}}{(\gamma/2) + 2\ln 2 + \frac{1}{2}\ln c} \tag{6}$$

for the near vertical rise portions of the λ_n curves shown on Fig. 2. Here, $\gamma = 0.5772156649 \cdots$ is the Euler-Mascheroni constant. The remarkable accuracy of this approximation is shown on Fig. 5. Here, for

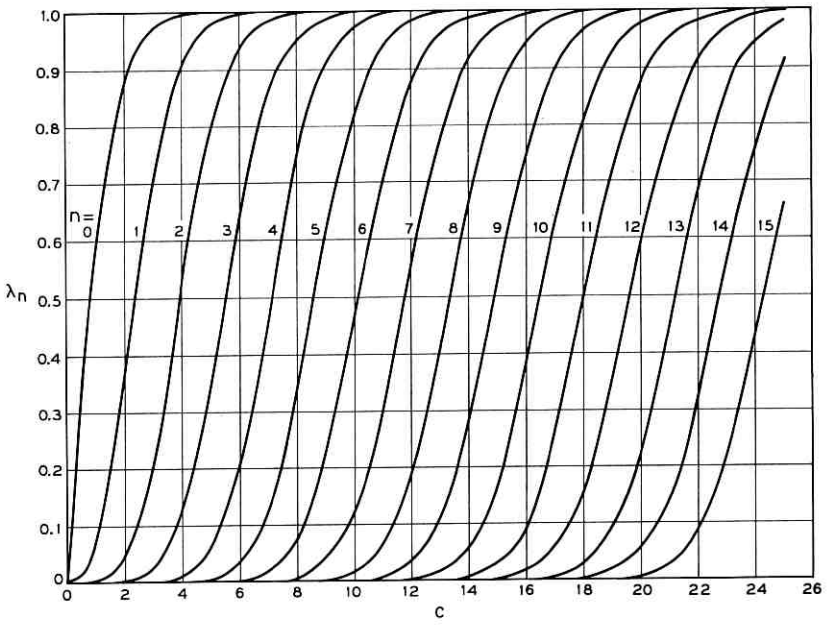


Fig. 2 — Eigenvalues, λ_n, of integral equation. Linear scale.

fixed values of n and $\hat{b}$, we have determined values of c from (6) and for these values of n and c have plotted $| (\hat{\lambda}_n/\lambda_n) - 1 |$ vs n. It is seen that for $0.2 \leqq \hat{\lambda}_n \leqq 0.9$, (5) and (6) give an excellent approximation even for small values of n.

Corresponding formulae for the χ_n follow. For fixed n and small c

$$\chi_n = n(n+1) + \frac{1}{2}\left[1 + \frac{1}{(2n-1)(2n+3)}\right]c^2 + 0(c^4)$$

and for fixed n and large c

$$\chi_n = (2n+1)c - \frac{2n^2+2n+3}{4} - \frac{(2n+1)(n^2+n+3)}{16c} + 0\left(\frac{1}{c^2}\right).$$

If n and c become large according to (3) with b fixed,

$$\chi_n = c^2 + 2bc + \frac{b^2-1}{2} - \frac{b^3-b}{8c} + 0\left(\frac{1}{c^2}\right).$$

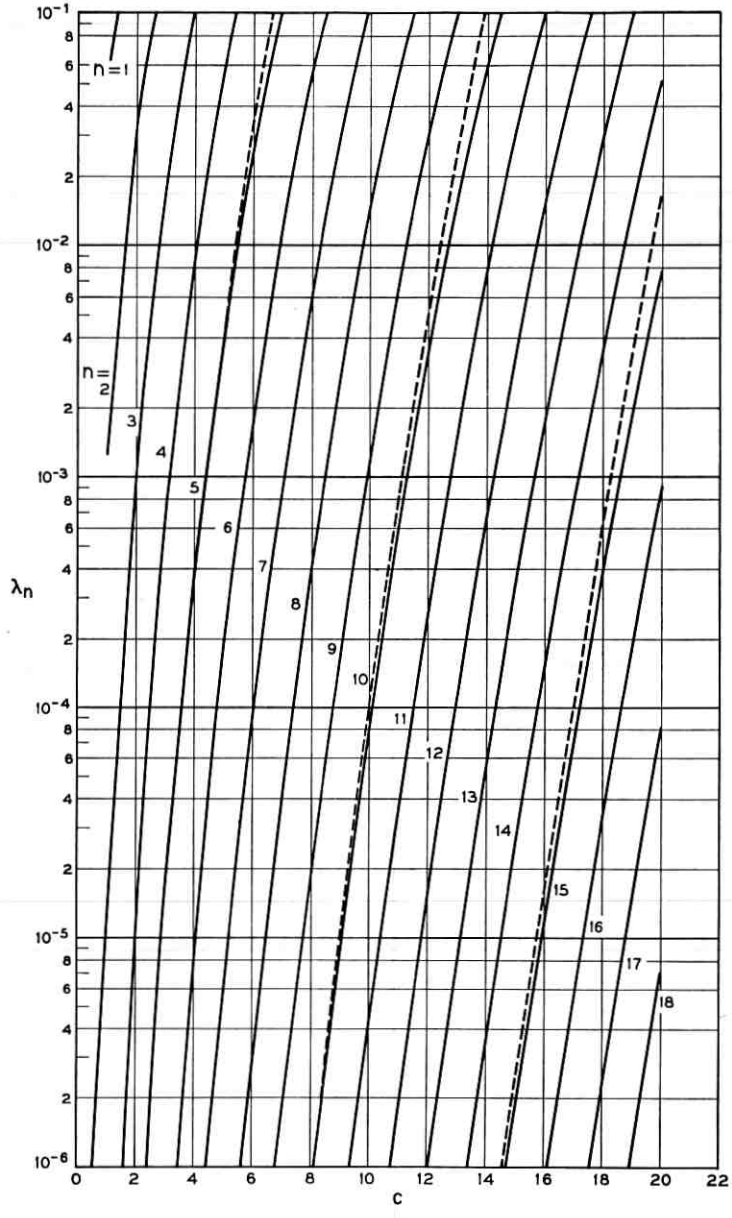


Fig. 3 — Eigenvalues, λ_n , of integral equation for $c < n\pi/2$. Dashed lines are approximation (1).

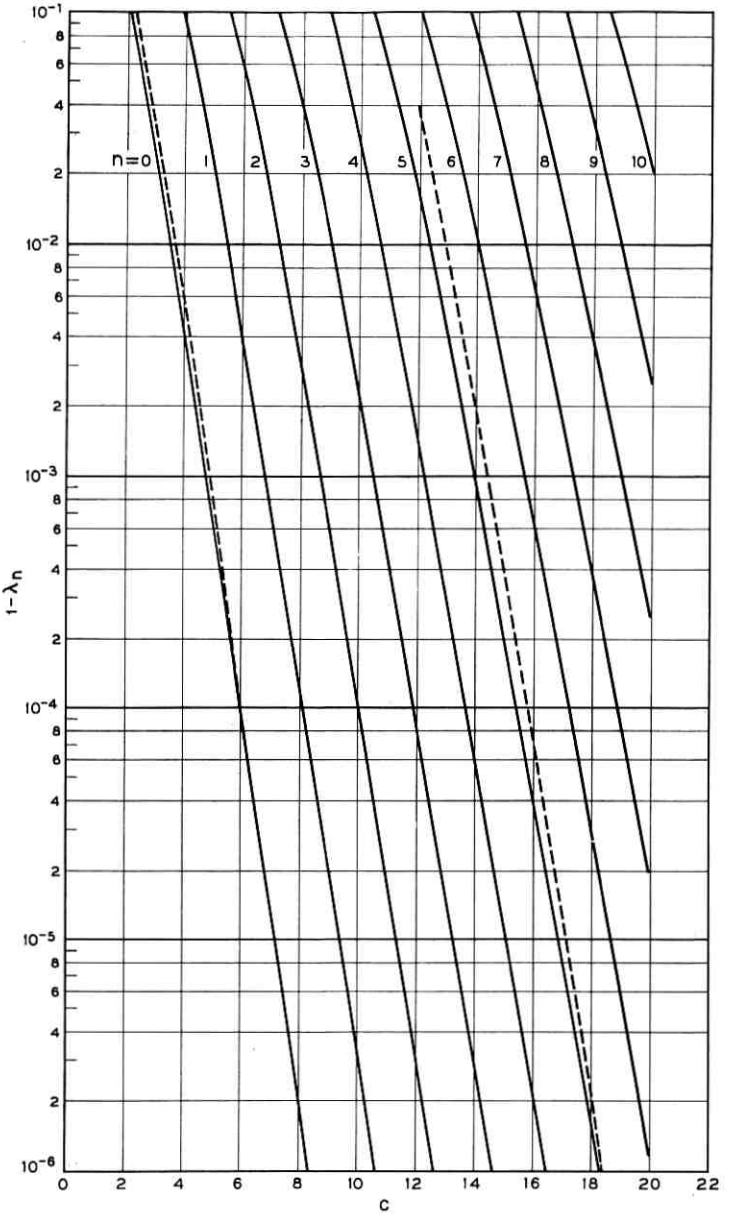


Fig. 4 — Eigenvalues, λ_n, of integral equation for $c > n\pi/2$. Dashed lines are approximation (2).

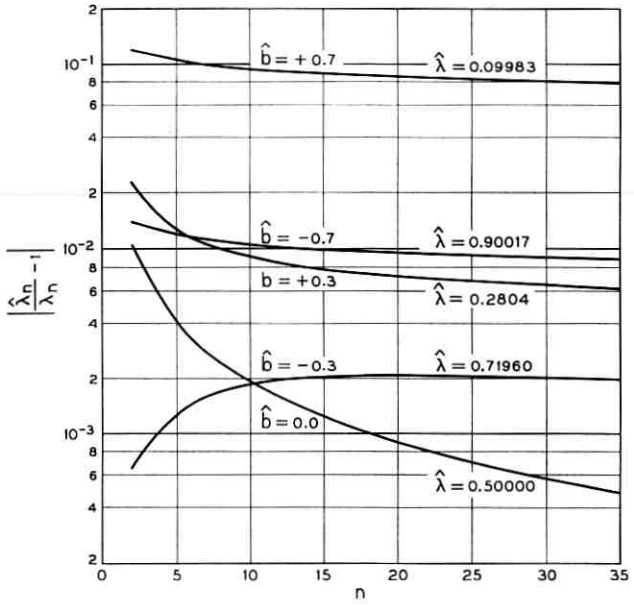


Fig. 5 — Accuracy of the approximation (5) — (6) for the eigenvalues λ_n .

REFERENCES

1. Flammer, C., *Spheroidal Wave Functions*, Stanford Univ. Press, Stanford, 1957.
2. Slepian, D., Some Asymptotic Expansions for Prolate Spheroidal Wave Functions, J. Math. and Phys., *44*, No. 2, June, 1965, pp. 99–140.

Cosine Sum Approximation and Synthesis of Array Antennas

By D. JAGERMAN

(Manuscript received May 10, 1965)

The problem of approximating a band-limited function, $H(t)$, by a sum of cosines arises in the design of phased array antennas. Three methods of synthesis are presented for establishing such designs. Error formulae are deduced for each method, including a new error formula for Tchebycheff quadrature. The existence of grating lobes is proved, and lower bounds for their location are developed.

I. INTRODUCTION

This paper is concerned with the problem of approximating a function $H(t)$ $(-\infty < t < \infty)$ by a cosine sum of the form

$$S_N(t) = \frac{1}{N}\sum_{j=1}^{N} \cos tx_j\,, \quad 0 \leqq x_1 < x_2 < \cdots < x_N \leqq 1. \tag{1}$$

The synthesis of array antennas is an application of the problem of this paper. Let isotropic radiating elements of strength $1/N$ be located along the x-axis at the points $x_j (1 \leqq j \leqq N)$ providing planar radiation of wavelength λ, and let θ designate the angle between the positive y-axis and a line passing through the origin and a far-field point, then, setting

$$t = \frac{2\pi \sin\theta}{\lambda}\,, \tag{2}$$

the far-field radiation pattern of the linear array is given by $S_N(t)$. The requirement that all the coefficients of the sum in (1) be equal generally stems out of the use of identical radiating elements, and out of the desire to employ identical feed for each element.

The function $H(t)$ represents the desired far-field radiation pattern; it will be required to satisfy the condition

$$H(t) = \int_0^1 F(x) \cos tx\, dx, \tag{3}$$

for some function $F(x) \in L(0,1)$, and the normalization condition

$$H(0) = 1. \tag{4}$$

When required, the function $F(x)$ will be extended to the interval $(-1,1)$ by

$$F(-x) = F(x). \tag{5}$$

The function $F(x)$ is, thus, the illumination required for a continuous aperture to produce the far-field pattern $H(t)$. Equation (3) defines the array aperture as one, and the function $H(t)$ to be bandlimited with bandwidth one.

The approximation or synthesis problem consists in the determination of the quantities $x_1, \cdots, x_N$ subject to the condition of (1) so that $S_N(t)$ shall approximate $H(t)$.

In this form, the problem is that of numerical quadrature by means of an equal-coefficient rule. Sections II and III present methods for accomplishing this. Section IV drops the restriction of equal coefficients and applies the well-known Gaussian quadrature rule. Section V discusses the existence of grating lobes and presents estimates for their location.

II. A RIEMANNIAN SUM METHOD

Let $H(t)$ be a characteristic function, that is, $H(t)$ satisfies the normalization condition (4) and the additional requirement

$$F(x) \geqq 0, \qquad 0 < x < 1, \tag{6}$$

then the function

$$L(x) = \int_0^x F(u)\, du \tag{7}$$

satisfies

$$L(0) = 0, \qquad L(1) = 1 \tag{8}$$

and is monotonic increasing. Let

$$y = L(x), \qquad x = G(y) \tag{9}$$

in which $G(y)$ is the function inverse to $L(x)$, then the required numbers x_j are given explicitly by

$$x_j = G\left(\frac{2j-1}{2N}\right). \qquad 1 \leqq j \leqq N. \tag{10}$$

The sum

$$S_N(t) = \frac{1}{N}\sum_{j=1}^{N} \cos\left[tG\left(\frac{2j-1}{2N}\right)\right] \tag{11}$$

is clearly a Riemannian sum for

$$I = \int_0^1 \cos\left[tG(y)\right] dy, \tag{12}$$

and hence

$$\lim_{N\to\infty} S_N(t) = I; \tag{13}$$

however,

$$I = \int_0^1 F(x) \cos tx \, dx = H(t) \tag{14}$$

and hence the approximation is secured. The error $R_N(t)$ given by

$$R_N(t) = H(t) - S_N(t) \tag{15}$$

will now be studied. For this purpose consider

Lemma 1:

$$c_n = \int_0^1 \sin\left[2\pi nL(x)\right] \sin tx \, dx$$

$$\Rightarrow R_N(t) = \frac{t}{\pi N}\sum_{k=1}^{\infty} (-1)^{k-1}\frac{c_{Nk}}{k}.$$

Proof: It will be convenient to introduce the function

$$S_N(t,y) = \frac{1}{N}\sum_{j=1}^{N} \cos\left[tG\left(y + \frac{2j-1}{2N}\right)\right]; \tag{16}$$

thus

$$S_N(t,0) = S_N(t). \tag{17}$$

The function $\cos\left[tG(y)\right]$ may be expanded into a Fourier series on the interval $(0,1)$; one has,

$$\cos\left[tG(y)\right] = H(t) + \sum_{n=1}^{\infty} a_n \cos 2\pi ny + \sum_{n=1}^{\infty} b_n \sin 2\pi ny, \tag{18}$$

in which

$$\begin{aligned} a_n &= 2\int_0^1 \cos\left[tG(y)\right] \cos 2\pi ny \, dy, \\ b_n &= 2\int_0^1 \cos\left[tG(y)\right] \sin 2\pi ny \, dy. \end{aligned} \tag{19}$$

Define c_n , d_n by

$$\begin{aligned} c_n &= \int_0^1 G'(y) \sin 2\pi ny \sin [tG(y)]\, dy \\ &= \int_0^1 \sin [2\pi nL(x)] \sin tx\, dx, \\ d_n &= \int_0^1 G'(y) \cos 2\pi ny \sin [tG(y)]\, dy \\ &= \int_0^1 \cos [2\pi nL(x)] \sin tx\, dx, \end{aligned} \tag{20}$$

then integration by parts applied to (19) yields

$$\begin{aligned} a_n &= \frac{t}{\pi n} c_n , \\ b_n &= \frac{1 - \cos t}{\pi n} - \frac{t}{\pi n} d_n . \end{aligned} \tag{21}$$

The Bernoullian function

$$\rho(y) = \tfrac{1}{2} - \{y\}, \tag{22}$$

in which $\{y\}$ designates the *fractional part* of y, has the Fourier series

$$\rho(y) = \sum_{n=1}^{\infty} \frac{\sin 2\pi ny}{\pi n}, \tag{23}$$

hence, replacing a_n , b_n in (18) by their values in (21), one obtains

$$\begin{aligned} \cos [tG(y)] = H(t) &+ \frac{t}{\pi} \sum_{n=1}^{\infty} \frac{c_n}{n} \cos 2\pi ny \\ &- \frac{t}{\pi} \sum_{n=1}^{\infty} \frac{d_n}{n} \sin 2\pi ny + (1 - \cos t)\rho(y). \end{aligned} \tag{24}$$

By summation of the geometric series, one has

$$\begin{aligned} \frac{1}{N} \sum_{j=1}^{N} \exp \left[i2\pi n \left(y + \frac{2j-1}{2N} \right) \right] &= e^{i2\pi ny}(-1)^{n/N}, \qquad N \mid n, \\ &= 0, \qquad N \nmid n, \end{aligned} \tag{25}$$

and hence, letting $n = Nk$ ($k > 0$ integral),

$$\frac{1}{N} \sum_{j=1}^{N} \cos 2\pi Nk \left(y + \frac{2j-1}{2N} \right) = (-1)^k \cos 2\pi Nky, \tag{26}$$

$$\frac{1}{N}\sum_{j=1}^{N}\sin 2\pi Nk\left(y+\frac{2j-1}{2N}\right)=(-1)^k\sin 2\pi Nky. \tag{27}$$

Equations (16), (24), (26), and (27) now yield

$$\begin{aligned}S_N(t,y) &= H(t)+\frac{t}{\pi N}\sum_{k=1}^{\infty}(-1)^k\frac{c_{Nk}}{k}\cos 2\pi Nky\\ &\quad-\frac{t}{\pi N}\sum_{k=1}^{\infty}(-1)^k\frac{d_{Nk}}{k}\sin 2\pi Nky\\ &\quad+(1-\cos t)\frac{1}{N}\sum_{j=1}^{N}\rho\left(y+\frac{2j-1}{2N}\right).\end{aligned} \tag{28}$$

The Fourier series for $\rho(y)$, (23), permits ready establishment of the identity

$$\sum_{j=1}^{N}\rho\left(y+\frac{2j-1}{2N}\right)=\rho\left(Ny+\frac{1}{2}\right), \tag{29}$$

hence

$$\begin{aligned}S_N(t,y) &= H(t)+\frac{t}{\pi N}\sum_{k=1}^{\infty}(-1)^k\frac{c_{Nk}}{k}\cos 2\pi Nky\\ &\quad-\frac{t}{\pi N}\sum_{k=1}^{\infty}(-1)^k\frac{d_{Nk}}{k}\sin 2\pi Nky\\ &\quad+\frac{1-\cos t}{N}\rho\left(Ny+\frac{1}{2}\right).\end{aligned} \tag{30}$$

Setting $y = 0$ in (30) yields the result of the lemma.

Lemma 2: $r \geqq 2$, *integral*, $W^{(r)}(x) \geqq \varepsilon_r > 0$ *or*

$$W^{(r)}(x) \leqq -\varepsilon_r < 0 \quad \textit{for} \quad a \leqq x \leqq b$$

$$\Rightarrow \left|\int_a^b \cos W(x)\,dx\right| \leqq r2^{(r+1)/2}\varepsilon_r^{-(1/r)}.$$

Proof: It is clear that only the inequality $W^{(r)}(x) \geqq \varepsilon_r > 0$ need be considered. The case $r = 2$ will be considered first. The function $W'(x)$ is monotonic increasing, hence it vanishes at most once in $[a,b]$, say at $x = c$, then

$$\int_a^b \cos W(x)\,dx = \int_a^c \cos W(x)\,dx + \int_c^b \cos W(x)\,dx. \tag{31}$$

Let $0 \leqq \delta \leqq b - c$ be chosen, then

$$\int_c^b \cos W(x)\, dx = \int_c^{c+\delta} \cos W(x)\, dx + \int_{c+\delta}^b \cos W(x)\, dx, \quad (32)$$

and hence

$$\left| \int_c^b \cos W(x)\, dx \right| \leqq \delta + \left| \int_{c+\delta}^b \cos W(x)\, dx \right|. \quad (33)$$

One has

$$\begin{aligned} \int_{c+\delta}^b \cos W(x)\, dx &= \int_{c+\delta}^b \frac{1}{W'(x)}\, d \sin W(x) \\ &= \frac{1}{W'(c+\delta)} \int_{c+\delta}^{\xi} d \sin W(x), \end{aligned} \quad (34)$$

in which the second mean-value theorem was used, and hence

$$\left| \int_{c+\delta}^b \cos W(x)\, dx \right| \leqq \frac{2}{W'(c+\delta)}. \quad (35)$$

Since

$$W'(c+\delta) = \int_c^{c+\delta} W''(x)\, dx \geqq \delta \varepsilon_2, \quad (36)$$

one obtains, from (33),

$$\left| \int_c^b \cos W(x)\, dx \right| \leqq \delta + \frac{2}{\delta \varepsilon_2}. \quad (37)$$

The choice

$$\delta = \sqrt{2}\, \varepsilon_2^{-\frac{1}{2}} \quad (38)$$

yields

$$\left| \int_c^b \cos W(x)\, dx \right| \leqq 2\sqrt{2}\, \varepsilon_2^{-\frac{1}{2}}. \quad (39)$$

The value of δ in (38) may exceed $b - c$, however, in this case the inequality of (39) is certainly correct since the integral always admits the estimate $b - c$.

Similarly choose $0 \leqq \delta \leqq c - a$, then

$$\int_a^c \cos W(x)\, dx = \int_a^{c-\delta} \cos W(x)\, dx + \int_{c-\delta}^c \cos W(x)\, dx, \quad (40)$$

and hence

$$\left|\int_a^c \cos W(x)\,dx\right| \leqq \left|\int_a^{c-\delta} \cos W(x)\,dx\right| + \delta. \tag{41}$$

One has

$$\begin{aligned}\int_a^{c-\delta} \cos W(x)\,dx &= \int_a^{c-\delta} \frac{1}{W'(x)}\,d \sin W(x) \\ &= \frac{1}{W'(c-\delta)} \int_\xi^{c-\delta} d \sin W(x),\end{aligned} \tag{42}$$

and hence

$$\left|\int_a^{c-\delta} \cos W(x)\,dx\right| \leqq -\frac{2}{W'(c-\delta)}. \tag{43}$$

Since

$$-W'(c-\delta) = \int_{c-\delta}^{c} W''(x)\,dx \geqq \delta\varepsilon_2. \tag{44}$$

one obtains from (41)

$$\left|\int_a^c \cos W(x)\,dx\right| \leqq \delta + \frac{2}{\delta\varepsilon_2}. \tag{45}$$

Hence

$$\left|\int_a^c \cos W(x)\,dx\right| \leqq 2\sqrt{2}\,\varepsilon_2^{-\frac{1}{2}}, \tag{46}$$

and, from (31),

$$\left|\int_a^b \cos W(x)\,dx\right| \leqq 4\sqrt{2}\,\varepsilon_2^{-\frac{1}{2}}. \tag{47}$$

The lemma is thus established for $r = 2$.

Induction will now be employed. The lemma is assumed true for $r = k \geqq 2$. Since $W^{(k+1)}(x) > 0$, $W^{(k)}(x)$ is monotonic increasing, and hence vanishes at most once in $[a,b]$, say at $x = c$. Choose $0 \leqq \delta \leqq b - c$, then

$$\int_c^b \cos W(x)\,dx = \int_c^{c+\delta} \cos W(x)\,dx + \int_{c+\delta}^b \cos W(x)\,dx, \tag{48}$$

and hence

$$\left|\int_c^b \cos W(x)\,dx\right| \leqq \delta + \left|\int_{c+\delta}^b \cos W(x)\,dx\right|. \tag{49}$$

The inductive hypothesis states

$$\left| \int_{c+\delta}^{b} \cos W(x)\, dx \right| \leqq k2^{(k+1)/2} W^{(k)}(c+\delta)^{-(1/k)}, \tag{50}$$

hence

$$\left| \int_{c}^{b} \cos W(x)\, dx \right| \leqq \delta + k2^{(k+1)/2} W^{(k)}(c+\delta)^{-(1/k)}. \tag{51}$$

Since

$$W^{(k)}(c+\delta) = \int_{c}^{c+\delta} W^{(k+1)}(x)\, dx \geqq \delta\varepsilon_{k+1}, \tag{52}$$

one has

$$\left| \int_{c}^{b} \cos W(x)\, dx \right| \leqq \delta + k2^{(k+1)/2} \delta^{-(1/k)} \varepsilon_{k+1}{}^{-(1/k)}. \tag{53}$$

The choice

$$\delta = 2^{k/2} \varepsilon_{k+1}{}^{-(1/k+1)} \tag{54}$$

yields

$$\left| \int_{c}^{b} \cos W(x)\, dx \right| \leqq (k+1)\, 2^{k/2} \varepsilon_{k+1}{}^{-(1/k+1)}. \tag{55}$$

The inequality of (55) remains correct even for $\delta > b - c$.

Similarly, choose $0 \leqq \delta \leqq c - a$, then

$$\int_{a}^{c} \cos W(x)\, dx = \int_{a}^{c-\delta} \cos W(x)\, dx + \int_{c-\delta}^{c} \cos W(x)\, dx, \tag{56}$$

and hence

$$\left| \int_{a}^{c} \cos W(x)\, dx \right| \leqq \left| \int_{a}^{c-\delta} \cos W(x)\, dx \right| + \delta. \tag{57}$$

The inductive hypothesis yields

$$\left| \int_{a}^{c} \cos W(x)\, dx \right| \leqq \delta + k2^{(k+1)/2} \left[-W^{(k)}(c-\delta) \right]^{-(1/k)}. \tag{58}$$

Since

$$-W^{(k)}(c-\delta) = \int_{c-\delta}^{c} W^{(k+1)}(x)\, dx \geqq \delta\varepsilon_{k+1}, \tag{59}$$

one has

$$\left| \int_a^c \cos W(x)\, dx \right| \leqq \delta + k2^{(k+1)/2}\delta^{-(1/k)}\varepsilon_{k+1}^{-(1/k)}. \tag{60}$$

Thus

$$\left| \int_a^c \cos W(x)\, dx \right| \leqq (k+1)\, 2^{k/2}\, \varepsilon_{k+1}^{-(1/k+1)}, \tag{61}$$

and hence

$$\left| \int_a^b \cos W(x)\, dx \right| \leqq (k+1)\, 2^{(k+2)/2}\, \varepsilon_{k+1}^{-(1/k+1)}. \tag{62}$$

The lemma is now established.

Theorem 1 provides an estimate of $R_N(t)$.

Theorem 1: $r \geqq 2$, *integral*, $L^{(r)}(x) \geqq \varepsilon_r > 0$ *or*

$$L^{(r)}(x) \leqq -\varepsilon_r < 0 \quad \textit{for} \quad 0 \leqq x \leqq 1$$

$$\Rightarrow |R_N(t)| \leqq r2^{(r+1)/2-(1/r)}\pi^{-1-(1/r)}\zeta(1+(1/r))\varepsilon_r^{-(1/r)}\,|t|\,N^{-1-(1/r)}.$$

Proof: One has, from Lemma 1,

$$c_n = \int_0^1 \sin[2\pi nL(x)] \sin tx\, dx, \tag{63}$$

and hence

$$c_n = \frac{1}{2}\int_0^1 \cos[2\pi nL(x) - tx]\, dx - \frac{1}{2}\int_0^1 \cos[2\pi nL(x) + tx]\, dx. \tag{64}$$

Lemma 2 applied to the integrals of (64) yields

$$|c_n| \leqq r2^{(r+1)/2-(1/r)}\pi^{-(1/r)}\varepsilon_r^{-(1/r)}n^{-(1/r)}. \tag{65}$$

The infinite series for $R_N(t)$ in Lemma 1 may now be estimated. Using (65), one obtains

$$|R_N(t)| \leqq r2^{(r+1)/2-(1/r)}\pi^{-1-(1/r)}\varepsilon_r^{-(1/r)}\,|t|\,N^{-1-(1/r)}\sum_{k=1}^{\infty}\frac{1}{k^{1+(1/r)}}. \tag{66}$$

Since the series of (66) is $\zeta(1+(1/r))$, the inequality of the theorem follows.

An example of the above analysis is provided by the choice

$$H(t) = J_0(t), \tag{67}$$

that is the Bessel function of first kind and order zero. For this case

$$F(x) = \frac{2}{\pi\sqrt{1 - x^2}}, \tag{68}$$

and hence

$$L(x) = \frac{2}{\pi} \sin^{-1} x. \tag{69}$$

Thus

$$x = \sin \frac{\pi}{2} y, \tag{70}$$

and

$$x_j = \sin \frac{\pi}{2} \frac{2j - 1}{2N}. \tag{71}$$

The function $L(x)$ satisfies

$$L'''(x) \geqq 1 = \varepsilon_3, \tag{72}$$

and hence, after numerical simplification, the error is estimated by

$$| R_N(t) | < 8 | t | N^{-(4/3)}. \tag{73}$$

Another example is given by

$$H(t) = (\sin \tfrac{1}{2}t/\tfrac{1}{2}t)^2. \tag{74}$$

One has

$$F(x) = 2 - 2x, \tag{75}$$

$$L(x) = 2x - x^2, \tag{76}$$

and

$$G(y) = 1 - \sqrt{1 - y}. \tag{77}$$

Thus

$$x_j = 1 - \sqrt{1 - (2j - 1)/2N}. \tag{78}$$

Since

$$L'' = -2 = -\varepsilon_2, \tag{79}$$

the error estimate obeys

$$| R_N(t) | < 1.3 | t | N^{-(3/2)}. \tag{80}$$

If $F(x)$ has high order of contact at the endpoints zero and one, then $H(t)$ will decrease rapidly with increasing t, and hence the sidelobes will be small. In particular, let

$$F^{(j)}(0) = 0, \qquad F^{(j)}(1) = 0, \qquad 0 \leqq j \leqq k \tag{81}$$

then, integration by parts applied to (3) yields

$$H(t) = -(-1)^{k/2} \frac{1}{t^{k+1}} \int_0^1 F^{(k+1)}(x) \sin tx \, dx, \qquad k \text{ even}, \tag{82}$$

and

$$H(t) = -(-1)^{(k-1)/2} \frac{1}{t^{k+1}} \int_0^1 F^{(k+1)}(x) \cos tx \, dx, \qquad k \text{ odd}. \tag{83}$$

If $F^{(k+1)}(x)$ is of bounded variation, then

$$| H(t) | \leqq \frac{V}{t^{k+2}}, \tag{84}$$

in which V is the total variation of $F^{(k+1)}(x)$. Equation (84) shows the rapid decay of the sidelobes.

An example of this type of tapered design is given by

$$F_k(x) = (2k + 1) \binom{2k}{k} [x(1 - x)]^k \tag{85}$$

which has order of contact $k - 1$ and for which

$$| H(t) | \leqq V/t^k. \tag{86}$$

In this case, V is the total variation of $F_k^{(k)}(x)$. Since

$$L^{(2k+1)}(x) = -(2k + 1)! \binom{2k}{k} = -\varepsilon_{2k+1} < 0 \tag{87}$$

one has, from Theorem 1,

$$| R_N(t) | \leqq E_k | t | N^{-1-(1/2k+1)}, \tag{88}$$

in which E_k is the constant determined by the theorem.

The function

$$H(t) = \sin t/t \tag{89}$$

corresponds to

$$F(x) = 1, \qquad L(x) = x, \qquad G(y) = y. \tag{90}$$

The distribution of radiators is

$$x_j = (2j - 1)/2N \tag{91}$$

and therefore is uniform. The approximability of this function is poor compared to the previous examples. Theorem 1 does not cover this case since $L''(x) = 0$; however, the Fourier coefficients c_n (20) may be explicitly evaluated, and the final determination of $R_N(t)$ obtained from Lemma 1. The result is

$$R_N(t) = \frac{t}{\pi N} \sum_{k=-\infty}^{\infty}{}' \frac{\sin (t + \pi k)}{k(t + 2\pi N k)}, \tag{92}$$

in which the prime shows the absence of the term $k = 0$. Evaluation of the integral

$$\int_0^1 \rho\left(Nx + \frac{1}{2}\right) \sin tx \, dx, \tag{93}$$

using (23), shows that

$$R_N(t) = -\frac{t}{N} \int_0^1 \rho\left(Nx + \frac{1}{2}\right) \sin tx \, dx. \tag{94}$$

Since

$$| \rho(Nx + \tfrac{1}{2}) | \leqq \tfrac{1}{2}, \qquad | \sin tx | \leqq 1, \tag{95}$$

one has

$$| R_N(t) | \leqq \tfrac{1}{2} | t | N^{-1}. \tag{96}$$

III. TCHEBYCHEFF QUADRATURE METHOD

Let $\varphi(x) \, \varepsilon \, C^{(M)}[-1,1]$, then the Tchebycheff quadrature formula[1] is

$$\int_{-1}^1 K(x)\varphi(x) \, dx \cong \frac{1}{M} \sum_{j=1}^{M} \varphi(\alpha_j), \qquad \int_{-1}^1 K(x) \, dx = 1. \tag{97}$$

The fundamental points α_j are determined by the conditions

$$M \int_{-1}^1 x^\nu K(x) \, dx = \sum_{j=1}^{M} \alpha_j{}^\nu = b_\nu, \qquad 0 \leqq \nu \leqq M. \tag{98}$$

Define the polynomial $\omega(z)$ by the polynomial portion of the Laurent expansion of

$$\exp\left(M \int_{-1}^1 K(x) \ln (z - x) \, dx\right) = z^M \exp\left(-\frac{b_1}{z} - \frac{b_2}{2z^2} - \frac{b_3}{3z^3} - \cdots\right) \tag{99}$$

about the origin,[2] then the zeros of $\omega(z)$ are the required numbers $\alpha_1, \cdots, \alpha_M$. This procedure yields an approximation which, by (98), is exact if $\varphi(x)$ is a polynomial of degree not exceeding M. To obtain an approximation to $H(t)$ of the required form (1), one may set

$$\varphi(x) = \cos tx, \qquad K(x) = \tfrac{1}{2}F(x), \qquad M = 2N; \tag{100}$$

the points $x_1, \cdots, x_N$ are now chosen as those α_j which are positive. Equations (3) and (97) yield the required result.

Define the error, ${R_N}^T$, of Tchebycheff quadrature by

$${R_M}^T = \int_{-1}^{1} K(x)\varphi(x)\,dx - \frac{1}{M}\sum_{j=1}^{M}\varphi(\alpha_j), \tag{101}$$

then Theorem 2 provides an estimate.

Theorem 2: The real numbers $\alpha_1, \cdots, \alpha_M$ are determined as the zeroes of the polynomial $\omega(x)$ defined in (99)

$$\Rightarrow \exists -1 < \xi < 1 \ni {R_M}^T = \frac{1}{M!}\int_{-1}^{1} K(x)\omega(x)\varphi^{(M)}(\xi)\,dx.$$

Proof: It will be shown that Tchebycheff quadrature is an instance of Newton-Cotes quadrature.

Define

$$\ell_j(x) = \frac{\omega(x)}{(x - \alpha_j)\omega'(\alpha_j)}, \tag{102}$$

then the Lagrange interpolation formula is

$$\varphi(x) = \sum_{j=1}^{M}\varphi(\alpha_j)\,\ell_j(x) + \frac{\varphi^{(M)}(\xi)}{M!}\,\omega(x), \tag{103}$$

in which ξ satisfies

$$\min\,(x, \alpha_1, \cdots, \alpha_M) < \xi < \max\,(x, \alpha_1, \cdots, \alpha_M). \tag{104}$$

The coefficients of the Newton-Cotes quadrature formula are given by

$$c_j = \int_{-1}^{1} K\,(x)\ell_j(x)\,dx \tag{105}$$

and hence, one has

$$\int_{-1}^{1} K(x)\varphi(x)\,dx = \sum_{j=1}^{M} c_j\varphi(\alpha_j) + \frac{1}{M!}\int_{-1}^{1} K(x)\omega(x)\varphi^{(M)}(\xi)\,dx. \tag{106}$$

Since Tchebycheff quadrature is exact when $\varphi(x)$ is a polynomial of

Thus, secondary lobes may be produced of strength nearly equal to the main beam. These are called *grating lobes*. Their existence is the subject of Theorem 5.

Theorem 5: There always exist grating lobes.

Proof: Dirichlet's theorem[4] on simultaneous approximation states:

Given $x_1, \cdots, x_N$, a positive integer q, and a positive integer τ_o, there exists a number τ in the range

$$\tau_o \leqq \tau \leqq \tau_o q^N, \tag{114}$$

and integers $p_1, \cdots, p_N$, such that

$$| \tau x_j - p_j | \leqq 1/q, \qquad 1 \leqq j \leqq N. \tag{115}$$

Accordingly, choose $\tau_o = 1$ and $t = 2\pi\tau$, then

$$tx_j = 2\pi\tau x_j = 2\pi p_j + (2\pi/q)\theta, \qquad | \theta | \leqq 1, \tag{116}$$

and

$$\cos tx_j = \cos \frac{2\pi}{q} \theta > 1 - (2\pi^2/q^2). \tag{117}$$

Thus

$$S_N(t) = \sum_{j=1}^{N} A_j \cos tx_j > 1 - \frac{2\pi^2}{q^2}. \tag{118}$$

Since q may be chosen arbitrarily large, the theorem is proved.

An inspection of all the error formulae of this paper shows the common feature that they increase with increasing $| t |$ and ultimately become trivial. For large $| t |$, $H(t)$ is small, hence, since $S_N(t)$, by Theorem 5, must ultimately become large, the error estimates must also become large. It follows that the grating lobe cannot occur until $R_N(t)$, ${R_N}^T(t)$, or ${R_N}^G(t)$ are at least one. The error estimates, therefore, provide a lower bound for the value of $| t |$ at which a grating lobe can occur. This is especially important in those designs where it is desired to eliminate the grating lobe from the scan sector. The methods of synthesis presented in this paper provide different estimates of location of the first grating lobe. Let T designate that location, then, for

$$H(t) = J_o(t), \qquad T > \tfrac{1}{8}N^{4/3}, \tag{119}$$

$$H(t) = (\sin \tfrac{1}{2}t/\tfrac{1}{2}t)^2, \qquad T > .77N^{3/2}, \tag{120}$$

$$H(t) = \sin t/t, \qquad T > 2N. \tag{121}$$

The above are the estimates obtained from the Riemannian sum method. The Tchebycheff and Gaussian quadrature methods do not yield estimates of grating lobe location nearly as advantageous as the Riemannian sum method. Thus, for

$$\begin{aligned} H(t) &= J_o(t), \qquad T > 4N, \\ H(t) &= \sin t/t, \qquad T > 4N. \end{aligned} \tag{122}$$

These results were obtained by rough approximations to the factorials in (112) and (113), however, they serve to show the difference between the Riemannian sum, and the Tchebycheff and Gaussian quadrature methods. Nonetheless, the last two methods may show a much smaller estimate of error for small $|t|$ than the Riemannian sum method.

REFERENCES

1. Tomlinson Fort, *Finite Differences*, Oxford University Press, 1948.
2. Milne-Thomson, L. M., *The Calculus of Finite Differences*, Macmillan and Co., 1933.
3. Sansone, G., *Orthogonal Functions*, Interscience, 1959.
4. Titchmarsh, E. C., *The Theory of the Riemann Zeta-Function*, Oxford University Press, 1951.

Energy Reception for Mobile Radio

By E. N. GILBERT

(Manuscript received July 2, 1965)

Statistical properties are derived for mathematical models of the multipath fading encountered in mobile radio. These properties are used to compare some receiving systems which use several antennas to combat fading. Particular attention is given to a system of J. R. Pierce which has electric and magnetic dipole antennas and computes the electromagnetic energy density at a point. The statistical properties considered here include energy density distribution functions, correlation coefficients, and the power spectrum of the energy density observed at a moving point.

I. INTRODUCTION

A radio signal may reach a receiver via several paths because of reflections from nearby objects. If the receiver is a mobile radio installation, the received field strength may fluctuate wildly because the reflected waves add with changing relative phases as the receiver moves. J. R. Pierce* has suggested a way to combat these fluctuations by using three antennas.

As background for Pierce's idea consider the standing wave pattern produced when a plane wave is reflected at normal incidence from a large wall. An electric dipole antenna moving toward the wall finds nulls in the electric field repeated at half-wavelength intervals. However, these nulls occur at maxima of the magnetic field. In fact, the total electromagnetic energy density $\frac{1}{2}(\epsilon \mid E \mid^2 + \mu \mid H \mid^2)$ is constant throughout the pattern.

In Pierce's scheme, the transmitter radiates a vertically polarized wave. The receiver carriers a vertical electric dipole antenna and also a pair of loop antennas with axes perpendicular to each other and to the dipole. These three antennas receive the three nonzero field components E_z, H_x, and H_y. The three antenna signals enter separate square-law detectors and the three detector outputs are added to

* Private communication.

obtain $\frac{1}{2}(\epsilon \mid E_z \mid^2 + \mu \mid H_x \mid^2 + \mu \mid H_y \mid^2) = \psi_T$, the total energy density. If the signal is amplitude modulated, the receiver may compute $\psi_T^{\frac{1}{2}}$ in order to achieve linear detection.

The output of this *energy receiver* remains constant as the receiver moves through the above-mentioned standing wave pattern near a wall. In more complicated interference patterns the total energy density does fluctuate, although hopefully not as much as the electric energy density alone. This paper examines some superpositions of vertically polarized plane waves in order to compare the energy receiver with a receiver which observes only the electric field. Two other receivers are examined briefly in Section V. One is a *diversity receiver* which has two or more electric dipole antennas and a switching system to select the antenna with the strongest signal. The second squares and adds the outputs of several electric dipoles.

Most of the analysis in this paper applies slightly more generally to a *weighted energy detector* which combines the electric and magnetic energy densities with weight factors $2d$ and $2b$ to obtain $d\epsilon \mid E \mid^2 + b\mu \mid H \mid^2$ (this is the energy density ψ_T if $d = b = \frac{1}{2}$). Sections III and IV show that certain unequal weights have some slight advantages.

In order to imitate the haphazardness of real mobile radio interference patterns most of the field models which follow assume waves with randomly chosen amplitudes, phases, and directions of propagation. Section IV finds probability distributions for the weighted energy density. At wavelengths longer than about 0.2 meters, such distributions might be used to predict the fraction of time that the signal will fade beyond the range of the receiver's AVC action. At shorter wavelengths, a fast automobile encounters fluctuations which have appreciable components at audible frequencies. Then questions about spectra (Section VIII) and correlations (Section VII) arise.

Fig. 1 shows the energy density as a function of position when four waves superimpose. The four waves had equal amplitudes but the phases and directions were chosen to typify some of the random models which follow. The propagation directions made angles of 0°, 60°, 140°, and 260° which were measured clockwise away from a horizontal direction. Table I gives the code for interpreting the printed symbols as energy densities in db above the mean level. The square in Fig. 1 is 3.6 wavelengths on a side. Fig. 2 uses the same four waves and the code of Table I to depict the electric energy density alone. It is immediately apparent that Fig. 2 represents a more violent function than Fig. 1. The peaks are higher (usually above 4 db), the valleys are deeper (often below −13 db), and intermediate levels are relatively scarce.

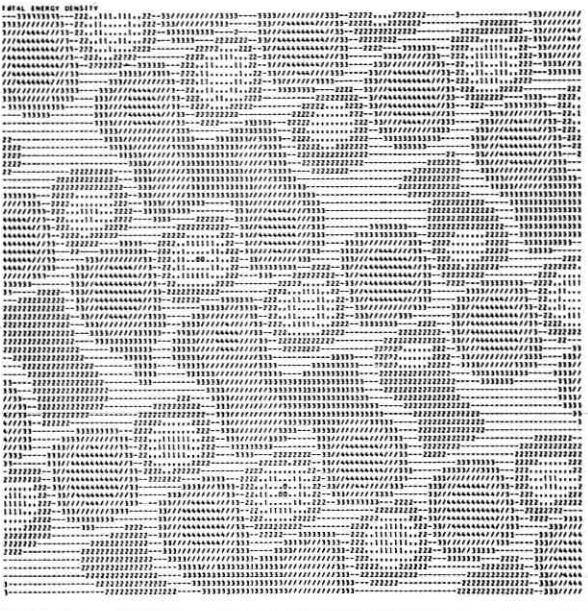


Fig. 1 — Total energy density of four superimposed waves.

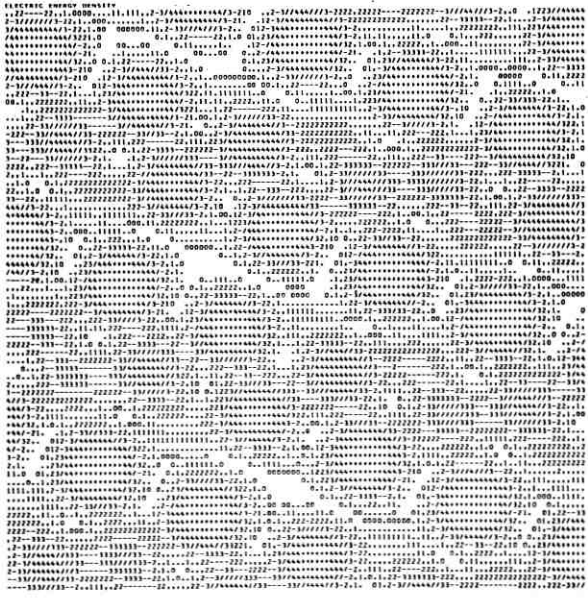


Fig. 2 — Electric energy density of four superimposed waves.

TABLE I — INTERPRETATION OF FIGS. 1 AND 2

Symbol	Interpretation
blank	below −13 db
0	between −13 db and −10 db
.	" −10 " " −7 "
1	" −7 " " −5 "
,	" −5 " " −3 "
2	" −3 " " −1 "
-	" −1 " " 0 "
3	" 0 " " 1 "
/	" 1 " " 2 "
4	" 2 " " 4 "
+	above 4 db

II. NOTATION

Fields will be functions of Cartesian coordinates (x,y) in a horizontal plane. The propagation direction of a vertically polarized wave will be specified by a unit vector $u = (u_x, u_y)$. Let P be the radius vector to a point in the (x,y) plane. At P, the following are the nonzero field components of a wave propagating in direction u:

$$E_z = \epsilon^{-\frac{1}{2}} A \exp\{-i\beta u \cdot P\}$$

$$H_x = \mu^{-\frac{1}{2}} u_y A \exp\{-i\beta u \cdot P\}$$

$$H_y = -\mu^{-\frac{1}{2}} u_x A \exp\{-i\beta u \cdot P\}$$

where $2\pi/\beta$ is the wavelength and A is a complex amplitude. All fields depend on time through a complex factor $\exp i\omega t$ which will be suppressed. The factors containing the dielectric constant ϵ and permeability μ were inserted to simplify the expressions for energy density. When waves from directions $u,v,w, \cdots$ are added, their amplitudes will be called $A(u), A(v), A(w), \cdots$. Most of this paper is concerned with the weighted energy density

$$\begin{aligned} \psi(P) = {} & d \left| \sum_u A(u) \exp - i\beta u \cdot P \right|^2 \\ & + b \left| \sum_u A(u) u_y \exp - i\beta u \cdot P \right|^2 \\ & + b \left| \sum_u A(u) u_x \exp - i\beta u \cdot P \right|^2. \end{aligned} \tag{1}$$

The coefficients d and b in (1) will always be nonnegative and will satisfy $d + b = 1$. Let $\psi_E(P)$ and $\psi_H(P)$ denote the electric and magnetic energy densities at P; then $\psi(P) = 2d\psi_E(P) + 2b\psi_H(P)$. The choice $d = b = \frac{1}{2}$ makes $\psi(P)$ the total energy density $\psi_T(P)$; $\psi(P)$ can also

become $2\psi_E(P)$ or $2\psi_H(P)$ if one adopts the extreme values $d = 1$ or $b = 1$. Another form of (1) is

$$\psi(P) = \sum_{u,v} A(u)A^*(v)(d + bu\cdot v) \exp i\beta(v - u)\cdot P \qquad (2)$$

where * denotes complex conjugate.

The average level ψ_0 about which $\psi(P)$ fluctuates, may be defined as the limit as $R \to \infty$ of the average of $\psi(P)$ over a circle of radius R. In the limit, terms of (2) with $v \neq u$ contribute zero to the average. Thus, if no two propagation directions are the same the average level is

$$\psi_0 = \sum_u |A(u)|^2. \qquad (3)$$

When $d = 1$ or $b = 1$, (3) shows that the average electric and magnetic energy densities are each $\frac{1}{2}\psi_0$.

According to (3), ψ_0 does not depend on d and b. The influence of d and b on some properties of the multipath fluctuations will be examined in subsequent sections, and (3) guarantees that the average detector output remains constant as d and b vary. When making such comparisons it must be recognized that the random noise received by the system may depend on d and b. For example, if the three antennas receive uncorrelated noises of equal power the noise output of the detector is proportional to $d + 2b$.

III. NULLS

Since three complex equations $E_z = 0$, $H_x = 0$, $H_y = 0$ must hold simultaneously at a point of zero energy density, it is not obvious when such a zero is possible. This section gives some examples of zeros.

When fewer than four waves add, no two having the same direction of propagation, no point can be a point of zero energy density. A proof of this fact is given in Appendix A. However, it is not necessarily desirable to have only a small number of waves. For example, consider two waves propagating in directions which differ by an angle ϑ, say $u = (\cos \frac{1}{2}\vartheta, \sin \frac{1}{2}\vartheta)$ and $v = (\cos \frac{1}{2}\vartheta, -\sin \frac{1}{2}\vartheta)$. Suppose the amplitudes are equal in magnitude but differ in phase by δ, say

$$A(u) = \exp(i(\varphi + \delta)), \qquad A(v) = \exp(i\varphi).$$

At a point $P = (x, y)$, (1) and (3) show that

$$\psi(P) = \psi_0\{1 + (d + b\cos\vartheta)\cos(\delta - 2\beta y \sin\tfrac{1}{2}\vartheta)\}. \qquad (4)$$

Note that $\psi(P)$ attains its minimum value $\psi_0\{1 - |d + b\cos\vartheta|\}$

along the family of lines

$$(2\beta \sin \tfrac{1}{2}\vartheta)y = \delta + \text{multiple of } \pi. \tag{5}$$

The "multiple" in (5) must be odd if $d + b \cos \vartheta > 0$ and even if $d + b \cos \vartheta < 0$. The (positive) minimum value can be arbitrarily small if ϑ is small enough.

It is interesting to compare the weighted energy density (4) with the electric energy density $\psi_E(P)$. When $d = 1$, (4) becomes

$$\psi_E(P) = \tfrac{1}{2}\psi_0\{1 + \cos(\delta - 2\beta y \sin \tfrac{1}{2}\vartheta)\} \tag{6}$$

a function which vanishes along the lines (5).

Equation (4) may be used to help decide a good choice of the coefficients d,b for a detector. Imagine the two waves produced at random in such a way that the angle ϑ and relative phase δ are independent random variables, both having probability density $(2\pi)^{-1}$ in the interval $(0, 2\pi)$. One wants $\psi(P)$ to fluctuate as little as possible. The variance of the random variable $\psi(P)$ is one measure of fluctuation. From (4) one obtains a variance

$$E\{|\psi(P) - \psi_0|^2\} = \tfrac{1}{2}\psi_0^2\{d^2 + \tfrac{1}{2}b^2\}$$

which has its minimum when $d = \frac{1}{3}$, $b = \frac{2}{3}$. Alternatively, one might prefer to pick d and b so that the expectation of the minimum value $1 - |d + b \cos \vartheta|$ of $\psi(P)$ is as large as possible. This condition requires d and b to minimize $E\{|d + b \cos \vartheta|\}$. The calculation given in Appendix B shows that the minimizing d and b are $d = 0.40$, $b = 0.60$. Although both criteria suggest that the magnetic energy density be weighted more than the electric energy density, both minima are so broad that an energy detector with $d = b = \frac{1}{2}$ does almost as well.

It is possible to have $\psi(0) = 0$ when waves from four different directions superimpose. For example, take the propagation directions along the $\pm x$, $\pm y$ coordinate axes

$$u = (1,0), \qquad v = (0,1), \qquad -u = (-1,0), \qquad -v = (0,-1)$$

and let the amplitudes be

$$A(u) = A(-u) = 1, \qquad A(v) = A(-v) = -1.$$

At the point $P = (x, y)$,

$$\psi(P) = \psi_0\{d(\cos \beta x - \cos \beta y)^2 + b(\sin^2 \beta x + \sin^2 \beta y)\}.$$

Zeros are spaced $\lambda/\sqrt{2}$ apart in this pattern if $b \neq 0$. However the

electric density is much worse,

$$\psi_E(P) = \tfrac{1}{2}\psi_0\{\cos \beta x - \cos \beta y\}^2.$$

The zeros of $\psi_E(P)$ are not isolated, they occupy two orthogonal families of parallel lines.

IV. ENERGY DISTRIBUTION FUNCTIONS

Three real parameters specify a wave, say the angle ϑ between u and the x axis, and the modulus $| A(u) |$ and phase of the complex amplitude $A(u)$. This section discusses some models which pick at random the $3N$ parameters of N interfering waves. In every case the N phases are chosen independently with constant probability density $(2\pi)^{-1}$ in the range $(0,2\pi)$. As a result, the models are stationary with respect to translations of the (x,y) coordinate system. In particular the probability distribution function of $\psi(P)$ is the same for all points P; to simplify the analysis take $P = 0$, the origin. The distribution function

$$F(\psi) = \text{Prob}\,\{\psi(0) \leqq \psi\}$$

is the probability that the detector of a mobile radio station produces an output less than ψ. In this section, the N waves have roughly the same statistical properties. By contrast Section VI considers a model in which one of the waves represents a strong wave direct from the transmitter.

In the first model the number of waves is $N \geqq 3$. The N propagation vectors $u_1, \cdots, u_N$ are not random. They are equally spaced around the unit circle; u_k makes angle $2\pi k/N$ with the positive x-axis. Each complex amplitude $A(u)$ will have the form $A(u) = R(u) + iI(u)$ where the $2N$ real numbers $R(u_1), \cdots, R(u_N), I(u_1), \cdots, I(u_N)$ are supposed independent Gaussian random numbers with mean 0 and variance 1. Another way to obtain the same random process is to pick moduli $| A(u_1) |, \cdots, | A(u_N) |$ independently from a Rayleigh distribution and the N phases independently with constant density $(2\pi)^{-1}$ in the range 0 to 2π.

According to (3), the average of $\psi(P)$ over the plane is

$$\psi_0 = \sum_u R^2(u) + \sum_u I^2(u)$$

so that the expected average is

$$\bar{\psi} = E(\psi_0) = 2N. \tag{7}$$

Appendix C derives the distribution function

$$F(\psi) = 1 - c^2 \exp(-\psi'/d) + (c-1)\{c + 1 + 2\psi'/d\} \exp(-2\psi'/b) \tag{8}$$

where $c = 2d/(2d - b)$ and $\psi' = \psi/\bar{\psi}$. Note that the number of waves N enters (8) only through the normalizing factor $\bar{\psi} = 2N$. The distribution function for the total energy is (8) with $c = 2$. From the limiting cases $d = 1$ and $d = 0$ of (8) one obtains distribution functions for the electric and magnetic energy densities.

$$\text{Prob}\{\psi_E(0) \leqq (\tfrac{1}{2}\bar{\psi})\psi'\} = 1 - \exp(-\psi')$$

$$\text{Prob}\{\psi_H(0) \leqq (\tfrac{1}{2}\bar{\psi})\psi'\} = 1 - (1 + 2\psi') \exp(-2\psi').$$

Curves EN and TN in Fig. 3 show the distributions of electric and total energy densities plotted in db above their respective mean average levels. The electric energy density has a much higher probability of being small than the total energy density. A curve for the magnetic energy density will be given in Section V.

The distribution obtained for $\psi_E(0)$ implies that the electric field strength $|E_z|$ has a Rayleigh distribution. In this respect, the model agrees with some experimental data of W. R. Young[5] (see in particular his Fig. 5).

For small values of ψ, (8) becomes

$$F(\psi) = (\tfrac{2}{3})d^{-1}b^{-2}\psi'^3 + \cdots$$

with missing terms $\cdots$ of order $0(\psi'^4)$ (see (24)). To make small values of ψ as unlikely as possible, one may minimize $d^{-1}b^{-2}$ by picking $d = \frac{1}{3}$, $b = \frac{2}{3}$. Recall that these values had another minimizing property in Section III. Again the advantage over using $d = b = \frac{1}{2}$ is slight. If the curve for $d = \frac{1}{3}$, $b = \frac{2}{3}$ were plotted in Fig. 3 it would lie about $\frac{1}{4}$ db to the right of the total energy density distribution curve. Equation (8) becomes indeterminate when $d = \frac{1}{3}$, $b = \frac{2}{3}$. However, in this special case, $\psi(0)$ has a chi-squared distribution of six degrees of freedom

$$F(\psi) = 1 - (1 + 3\psi' + \tfrac{1}{2}(3\psi')^2) \exp - 3\psi'.$$

Part of the variability of $\psi(0)$ comes from the randomness of the average value ψ_0. For any particular choice of the wave amplitudes, ψ_0 will not be exactly $\bar{\psi}$; the distribution of $\psi(0)/\psi_0$ might have been more relevant. However, ψ_0 has a chi-squared distribution with $2N$ degrees of freedom and so has high probability of being close to $\bar{\psi}$, say within 0.5 db, especially when N is large.

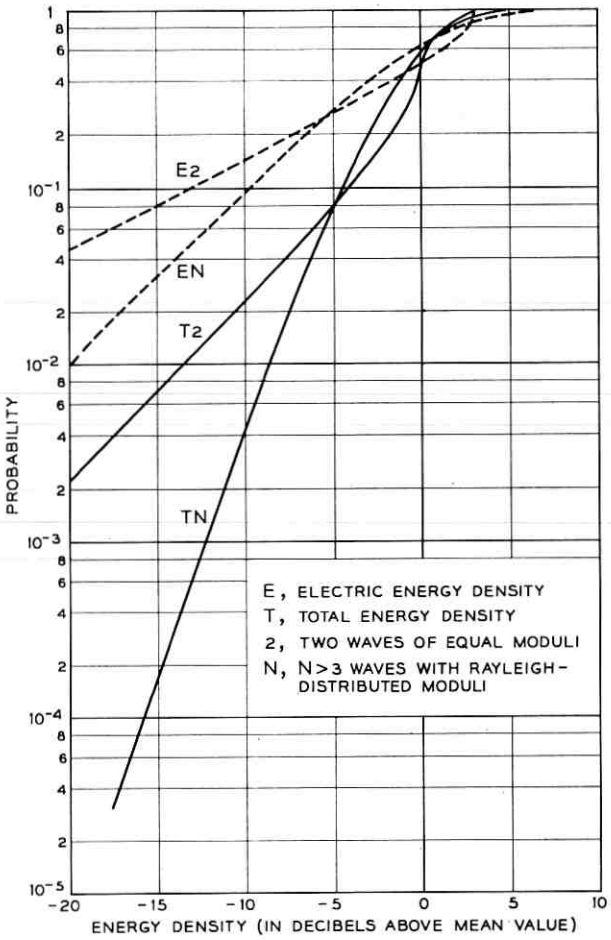


Fig. 3 — Probability distribution functions for energy densities.

The simple form of the distribution (8) results from the special random process which picks the amplitudes and directions. One might prefer to choose directions independently at random with probability $\vartheta/2\pi$ of making an angle less than ϑ with the x-axis. Other amplitude distributions also suggest themselves. It seems reasonable to continue to insist on independent amplitudes with random phases but one might use equal moduli $| A(u) |$ or another modulus distribution instead of the Rayleigh distribution. Undoubtedly $F(\psi)$ will be a more complicated function of N in these cases. However, (8) must still apply in the limit

of large N. For example, let M be an integer approximately equal to $N^{\frac{1}{2}}$. Let $v_1, \cdots, v_M$ be M unit vectors equally spaced around the unit circle. For each wave direction u find the vector u' in the list $v_1, v_2, \cdots, v_M$ which approximates u as closely as possible (and so to within angle π/M). If N is large, one makes only a small error in $\psi(0)$ by replacing each true direction u in (1) by its approximation u'. This approximation replaces the waves with random directions by waves with M equally spaced directions. For $i = 1, \cdots, M$ the approximating waves with direction v_i add up to a single wave, say with amplitude $A'(v_i)$. The central limit theorem shows that the real and imaginary parts of $A'(v_i)$ have approximately Gaussian distributions when N is large. Then the assumptions leading to (8) hold again.

Limiting results may be misleading if applied when the number of waves is small, as may be typical in mobile radio. Curves $T2$ and $E2$ of Fig. 3 show distributions for the total energy density and electric energy density (plotted in db above their mean values) for a superposition of two waves with equal moduli, random phases, and random directions. The distribution of total energy density $\psi_T(0)$ was obtained numerically using (4) with $d = b = \frac{1}{2}$; $\psi_T(0)$ was evaluated for 200 equally spaced values of ϑ and 200 equally spaced values of δ. A histogram of the 40,000 numbers was compiled to get the distribution. The electric energy distribution is easily derived from (6):

$$\text{Prob}\,\{\psi_E(0) \leqq (\tfrac{1}{2}\psi_0)\psi'\} = \pi^{-1} \arccos\,(1 - \psi')$$

(see Margaret Slack[4] for the electric energy distribution when other numbers of random waves of equal moduli combine). The curves show that the case of two waves of equal moduli is much worse for mobile radio than the case of waves with Rayleigh distributed moduli. Nevertheless total energy detection is again much better than electric energy detection.

When more than two waves of equal moduli, random phases, and random directions combine, the total energy density $\psi_T(0)$ depends on many random parameters. To find its distribution function by a numerical integration of the kind used for two waves would be much too costly. A computer experiment was used instead. Using pseudo-random numbers to pick phases and directions, the computer generated a sequence of field components for independent plane waves. After computing each new wave the computer found for $N = 2, \cdots, 10$, the energy density in the sum of the N most recently computed waves. After 10,009 waves, the computer had compiled histograms of the energy

densities in sums of 2, 3, $\cdots$, 10 waves, each based on 10,000 random samples. The same wave appeared in N consecutive sums of N waves; then samples closer together than N were not independent. However, the estimate of $F(\psi)$ is at least as good as if the experiment had $10{,}000/N$ independent samples.

Table II summarizes this experiment and compares the observations with theoretical predictions based on the curves in Fig. 3. The numbers observed agree surprisingly well with (8) even when $N = 3$. A comparison of the theoretical and observed numbers for $N = 2$ gives an idea of the accuracy of the experiment.

V. RECEPTION USING M ELECTRIC DIPOLES

Let m vertical dipole antennas be placed at points $P_1, \cdots, P_m$. Using switched diversity reception the received signal energy density is

$$\psi_D = \max \{\psi_E(P_1), \cdots, \psi_E(P_m)\}.$$

Another possibility is to square the antenna signals and add them (additive diversity); then the detector output is

$$\psi_S = \psi_E(P_1) + \psi_E(P_2) + \cdots + \psi_E(P_m).$$

$P_1, \cdots, P_m$ will be assumed spaced so far apart that the fields at these points may be considered independent random variables. If there are N waves generated at random by the first model of Section III, then each term $\psi_E(P_i)$ is a random variable with the chi-squared distribution of two degrees of freedom and mean N.

TABLE II — FRACTION OF SUMS OF N WAVES OF ENERGY DENSITY $\leqq \psi$. SAMPLE SIZE IS 10,000

ψ/ψ_0 in db	$N = 2$ theor	$N = 2$ obs	$N = 3$ obs	$N = 5$ obs	$N = 10$ obs	large N Eq. (8)
−16	0.0057	0.0066	0.0002	0	0	0.00008
−14	0.0091	0.0084	0.0002	0.0003	0	0.0003
−12	0.0144	0.014	0.0004	0.0003	0.001	0.0011
−10	0.0232	0.025	0.0004	0.002	0.003	0.0042
−8	0.0376	0.038	0.006	0.009	0.010	0.0144
−6	0.0614	0.063	0.030	0.033	0.047	0.0459
−4	0.1037	0.104	0.105	0.120	0.117	0.1301
−2	0.1851	0.185	0.227	0.270	0.296	0.3103
0	0.5000	0.500	0.530	0.574	0.584	0.5869
2	0.8908	0.891	0.882	0.853	0.849	0.8484
4	1	1	0.995	0.989	0.979	0.9742
6	1	1	1	1	1	0.9986

The distribution function for ψ_D is

$$F_D(\psi) = \text{Prob } (\psi_D \leqq \psi) = \prod \text{Prob } \{\psi_E(P_i) \leqq \psi\}$$
$$= \{1 - \exp (-\psi/N)\}^m.$$

The mean of ψ_D is

$$\bar{\psi}_D = \int_0^\infty \{1 - F_D(\psi)\} d\psi = \sum_{k=1}^{m} \binom{m}{k} (-1)^{k+1}/k.$$

For $m = 1, 2, \cdots, 5$, $\bar{\psi}_D$ is N, $3N/2$, $11N/6$, $25N/12$, $137N/60$. Then

$$F_D(\psi) = \{1 - \exp (-3\psi/2\bar{\psi}_D)\}^2 \qquad \text{when } m = 2$$
$$F_D(\psi) = \{1 - \exp (-11\psi/6\bar{\psi}_D)\}^3 \qquad \text{when } m = 3, \text{ etc.}$$

ψ_S has the chi-squared distribution with $2m$ degrees of freedom and mean Nm. Fig. 4 shows the distribution function of ψ_S with $m = 2$, 3, 5, and 8. The curves for the distributions of ψ_D and ψ_S lie very close and so the curves for ψ_D were not added to Fig. 4.

When $m = 3$ the distribution function of ψ_S is exactly the same as the one for the weighted energy density $\psi(0)$ with $d = \frac{1}{3}$ and $b = \frac{2}{3}$. As noted in Section IV, this distribution function is slightly better than the one for the energy density ($d = b = \frac{1}{2}$) at small values of ψ.

In Section IV, the two magnetic field components $H_x(0)$ and H_y were found to be uncorrelated. Then the distribution function of the magnetic energy density at zero follows the curve labeled $m = 2$ in Fig. 4.

VI. STRONG DIRECT WAVE

Section IV presented extreme cases in which the waves are all roughly of comparable strength. In this section another wave, stronger than the others, will be added to represent a "direct" wave of amplitude R. It is no longer easy to derive $F(\psi)$ exactly. However, the asymptotic form of $F(\psi)$ for small values of ψ is derived in Appendix C using the first model of Section IV. This result (25) assumes a convenient form in terms of the quantity

$$\sigma = 2N/\bar{\psi} = 1 - R^2/\bar{\psi}$$

which represents the fraction of the expected weighted energy density $\bar{\psi}$ contributed by the N scattered waves. When N and R are expressed in terms of $\bar{\psi}$ and σ the final result is

$$F(\psi) = (\tfrac{2}{3}) d^{-1} b^{-2} \sigma^{-3} (\psi/\bar{\psi})^3 \exp - 3(\sigma^{-1} - 1) \qquad (9)$$

approximately for small ψ.

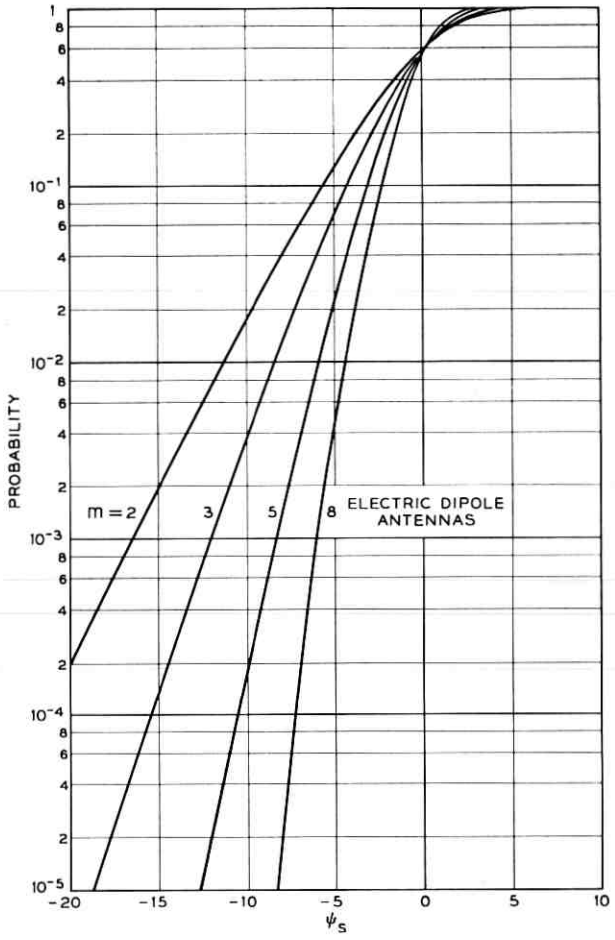


Fig. 4 — Probability distribution function for the sum ψ_s of the electric energy densities at m points.

The case $R = 0$, or $\sigma = 1$, was discussed in Section IV. When $0 < \sigma < 1$, (9) contains the extra factor

$$\sigma^{-3} \exp - 3(\sigma^{-1} - 1)$$

which is less than 1 and approaches 0 with decreasing σ. One concludes that $F(\psi)$ is then smaller than the value given by (8); i.e., deep minima tend to be less frequent when a direct wave is present. For example, if scattered waves account for only half the received weighted energy density ($\sigma = \frac{1}{2}$), the probability of receiving less weighted energy

density than ψ is only 8 exp $-$ 3 = 0.40 times the probability (8) for $\sigma = 1$.

VII. CORRELATION COEFFICIENTS

This section finds correlation coefficients between various pairs of energy densities. The correlation coefficient between $\psi(0)$ and $\psi(P)$ indicates whether a receiver traveling from 0 to P will find very different weighted energy densities at the two points. The correlation coefficient between $\psi_E(0)$ and $\psi_E(P)$ might be used to decide whether 0 and P are good locations for two electric dipole antennas in a diversity system; one would want low or negative correlation. For similar reasons, correlation coefficients involving the magnetic energy density $\psi_H(P) = \psi(P) - \psi_E(P)$ are interesting.

The waves in this section are produced by a random process slightly more general than the one used to get Table I. The N propagation directions and N phases are chosen at random and independently as in Section IV. The moduli $| A(u_1(| , \cdots , | A(u_N) |$ are now chosen independently from a common probability distribution. It is not necessary to know the distribution in detail. Only the expected values of the second and fourth powers $E(| A |^2)$, $E(| A |^4)$ enter into the correlation coefficients. To facilitate comparisons with Section IV, take $E(| A |^2) = 2$. Then the expected average weighted energy density is again

$$\bar{\psi} = 2N.$$

R. H. Clarke has also used this random process in an unpublished study of some different correlations.

It will be convenient to express the fourth power moment as

$$E(| A |^4) = 4 + \Sigma^2,$$

so that Σ^2 is the common variance of the squared moduli. When the moduli are all the same, $| A(u_i) | = 2^{\frac{1}{2}}$, $i = 1, \cdots , N$, and the variance Σ^2 is zero. When moduli have the Rayleigh distribution, $\Sigma^2 = 4$.

All the correlation coefficients of interest will be obtained as special cases of a single result. Consider two weighted energy densities.

$$\psi_1 = 2d\psi_E(0) + 2b\psi_H(0)$$

and

$$\psi_2 = 2D\psi_E(P) + 2B\psi_H(P).$$

Appendix D proves

$$E(\psi_1\psi_2) - E(\psi_1)E(\psi_2) = N\Sigma^2 + 4N(N-1)\{dDJ_0^2(\beta r) + (dB + bD)J_1^2(\beta r) + \tfrac{1}{2}bB(J_0^2(\beta r) + J_2^2(\beta r))\} \quad (10)$$

where $r = |P|$. When $D = d$, $B = b$, and $P = 0$, then $\psi_2 = \psi_1$ and (10) becomes the variance of ψ_1

$$\text{Var } \psi_1 = N\Sigma^2 + (4d^2 + 2b^2)N(N-1). \quad (11)$$

Likewise,

$$\text{Var } \psi_2 = N\Sigma^2 + (4D^2 + 2B^2)N(N-1). \quad (12)$$

The coefficient of correlation between ψ_1 and ψ_2 is

$$\rho = \{E(\psi_1\psi_2) - E(\psi_1)E(\psi_2)\} / \{\text{Var } \psi_1 \text{ Var } \psi_2\}^{\frac{1}{2}}, \quad (13)$$

which may be evaluated using (10), (11), and (12). The case of equal moduli ($\Sigma^2 = 0$) is especially simple because then (assuming $N > 1$) the factors $N(N-1)$ cancel out and ρ does not depend on N. This is the only case in which $\rho \to 0$ as $r \to \infty$; when $\Sigma^2 > 0$ the average weighted energy density ψ_0 is uncertain and so the energy densities remain slightly correlated even at points far apart. When N is large this residual correlation is small and ρ approaches its value for equal moduli.

By choosing special values for d, b, D, and B, one can obtain correlation coefficients of special interest:

$$\rho_{TT} = \{3J_0^2(\beta r) + 4J_1^2(\beta r) + J_2^2(\beta r)\}/3$$

$$\rho_{EE} = J_0^2(\beta r)$$

$$\rho_{HH} = J_0^2(\beta r) + J_2^2(\beta r)$$

$$\rho_{ET} = (\tfrac{2}{3})^{\frac{1}{2}}\{J_0^2(\beta r) + J_1^2(\beta r)\}$$

$$\rho_{EH} = 2^{\frac{1}{2}}J_1^2(\beta r)$$

$$\rho_{HT} = 3^{-\frac{1}{2}}\{J_0^2(\beta r) + 2J_1^2(\beta r) + J_2^2(\beta r)\}.$$

Here the subscripts E,H,T indicate the kind of energy, electric, magnetic, or total, at the two points. For the sake of simplicity the coefficients have been given only in the special case of equal moduli ($\Sigma^2 = 0$) or in the limit of large N. Some of the more interesting coefficients are plotted in Fig. 5.

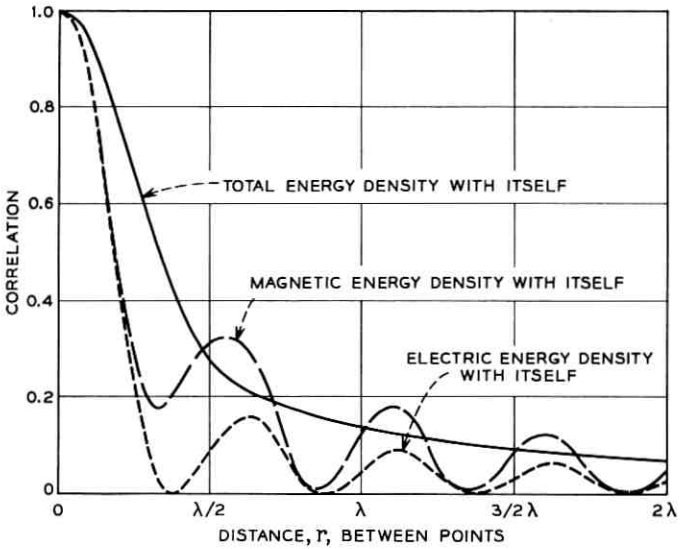


Fig. 5 — Coefficient of correlation between the energy densities at two points separated by distance r.

VIII. SPECTRA

When a receiver moves with constant velocity vector V_0 the received energy density $\psi(V_0t)$ is a random function of time. If one assumes the model of Section VII, the autocorrelation function of $\psi(V_0t)$ is known and hence its power spectrum may be found. The power spectrum of $\psi(V_0t)$ gives some idea of the frequencies at which the fluctuation noise is likely to be strong. In particular cases, depending on the way that the modulating signal is to be extracted from $\psi(P)$, the spectra of other functions may be more important. For example, if an AM system is used, one might prefer to know the power spectrum of $\psi^{\frac{1}{2}}(V_0t)$. For another kind of fading spectrum see J. F. Ossanna[3]. Ossanna combines two random waves and derives the spectrum obtained at the output of an envelope detector receiving the electric field.

For purposes of comparing power spectra it is convenient to normalize them to make the total power unity. Appendix E takes the Fourier transform of $E\{\psi(0)\psi(V_0t)\}/E\{\psi^2(0)\}$ to obtain a normalized spectrum. In order to keep formulas simple, Appendix E and this section consider only the case of large N.

The normalized power spectrum of $\psi(V_0t)$ contains a spectral line at zero frequency which represents the carrier or desired signal. The power

in this line is $1/(1 + d^2 + \frac{1}{2}b^2)$. Again the choice $d = \frac{1}{3}$, $b = \frac{2}{3}$ maximizes this power. The rest of the spectrum is fluctuation noise distributed with a spectral density function $s(f)$. Fig. 6 shows this spectrum for the total energy density and for the electric energy density. Both spectra vanish when f is larger than a cutoff frequency

$$f_0 = 2 \mid V_0 \mid /\lambda. \tag{14}$$

Note that f_0 is the frequency of the fluctuations in electric energy density observed by a vehicle moving toward the wall in the interference pattern described in Section I. When $0 < f \leqq f_0$, the spectral density has an analytic expression (32) in terms of the complete elliptic integrals $K(x)$ and $E(x)$. Let $\nu = f/f_0$. Then

$$s(f) = \frac{16}{33\pi^2 f_0} \{(3 - \nu^2)K((1 - \nu^2)^{\frac{1}{2}}) - 2(2 - \nu^2)E((1 - \nu^2)^{\frac{1}{2}})\} \tag{15}$$

for the total energy density and

$$s(f) = K\{(1 - \nu^2)^{\frac{1}{2}}\}/(\pi^2 f_0) \tag{16}$$

for the electric energy density.

The fluctuation noise $\psi_T(V_0 t)$ appears to be less troublesome than $\psi_E(V_0 t)$ both because it contains less total power away from the carrier

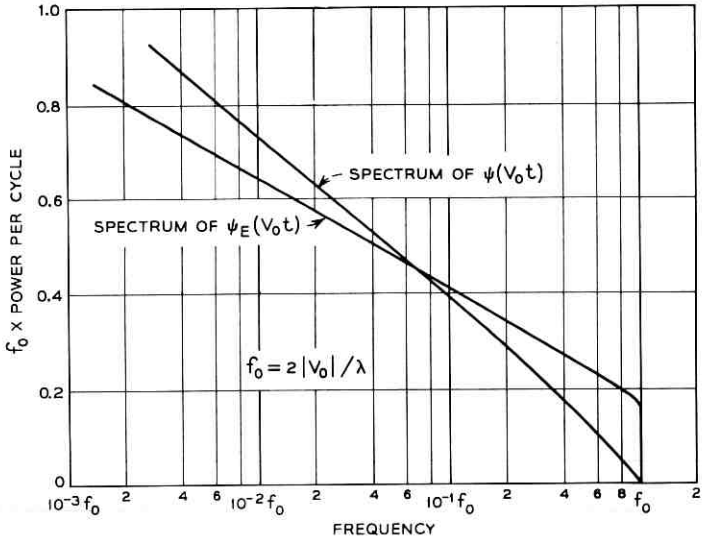


Fig. 6 — Power spectra of the energy densities $\psi(V_0 t)$ and $\psi_E(V_0 t)$ observed by a vehicle moving with constant velocity V_0.

and also because its spectrum is concentrated more toward low frequencies. As Fig. 6 shows, the spectrum of $\psi_T(V_0 t)$ goes to 0 smoothly at $f = f_0$ while the spectrum of $\psi_E(V_0 t)$ remains at a high level until it drops to 0 discontinuously at $f = f_0$.

IX. ACKNOWLEDGMENTS

I am grateful to J. R. Pierce, who first suggested this problem, R. H. Clarke, C. C. Cutler, and W. C. Jakes, Jr., for much helpful advice and orientation on the subject of mobile radio.

APPENDIX A

The impossibility of producing a zero by adding fewer than four waves

In what follows, waves must have nonzero amplitude and no two waves may have the same propagation direction. It is clearly possible for two or three waves with same direction to cancel if their amplitudes add up to zero.

The condition for a zero at the origin ($P = 0$) is

$$0 = \psi(0) = \left| \sum_u A(u) \right|^2 + \left| \sum_u A(u) u_y \right|^2 + \left| \sum_u A(u) u_x \right|^2$$

or simply

$$0 = \sum_u A(u) = \sum_u A(u) u. \tag{17}$$

Consider first the case of two waves in different directions u and v. Then, (17) becomes

$$A(u) + A(v) = 0.$$

$$A(u)u + A(v)v = 0.$$

The second (vector) equation requires that the unit vectors u,v be colinear. Since v cannot equal u, $v = -u$. Then

$$A(u) + A(v) = 0$$

$$A(u) - A(v) = 0,$$

a system with no solution except the trivial one $A(u) = A(v) = 0$.

When there are three waves with directions u,v,w, one may eliminate $A(w)$ from the system (17) to get

$$A(u)(u - w) + A(v)(v - w) = 0.$$

$A(u)$ cannot be zero. Then

$$u = w + a(v - w)$$

where $a = -A(v)/A(u)$. Since $|u|^2 = |w|^2 = 1$, one finds

$$0 = 2aw \cdot (v - w) + a^2 |v - w|^2. \tag{18}$$

Also,

$$0 = 2w \cdot (v - w) + |v - w|^2 \tag{19}$$

follows similarly from $v = w + (v - w)$ and $|v|^2 = |w|^2 = 1$. Use (19) eliminate $2w \cdot (v - w)$ from (18) and get

$$|v - w|^2 a(a - 1) = 0.$$

Since $v \neq w$ and since $a \neq 0$ (otherwise $A(v) = 0$), $a = 1$. However, if $a = 1$, $u = w + 1(v - w) = v$, a contradiction.

APPENDIX B

Weights which maximize the expected minimum value of $\psi(P)$

In the interference pattern (4) for two random waves, $\psi(P)$ attains a minimum value

$$\psi_{\min} = \psi_0\{1 - |d + b \cos \vartheta|\}.$$

The expected value of $\psi_{\min}$ is

$$E(\psi_{\min}) = \psi_0 \left\{1 - (2\pi)^{-1} \int_0^{2\pi} |d + b \cos \vartheta| \, d\vartheta\right\}. \tag{20}$$

The evaluation of the integral in (20) requires two cases. First, if $b \leqq \frac{1}{2} \leqq d$, $|d + b \cos \vartheta| = d + b \cos \vartheta$ and

$$E(\psi_{\min}) = \psi_0 b, \qquad (b \leqq d). \tag{21}$$

Second, if $d \leqq \frac{1}{2} \leqq b$, let $\vartheta_0 = \cos^{-1}(d/b)$. Then

$$|d + b \cos \vartheta| = \begin{cases} d + b \cos \vartheta & \text{when } |\vartheta \leqq \pi - \vartheta_0 \\ -d - b \cos \vartheta & \text{otherwise} \end{cases}$$

and

$$E(\psi_{\min}) = \psi_0\{b + (2/\pi)(d\vartheta_0 - \sqrt{b^2 - d^2})\}, \qquad (d \leqq b). \tag{22}$$

Fig. 7 shows $E(\psi_{\min})$ plotted vs d. There is a broad maximum near $d = 0.4$, $b = 0.6$.

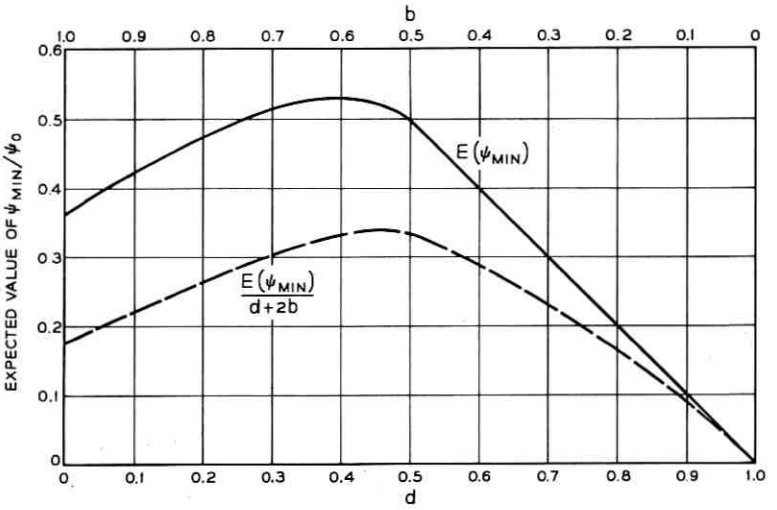


Fig. 7 — $E(\psi_{min})$ and $E(\psi_{min})/(d + 2b)$.

As mentioned in Section II, the received noise power will depend on d and b. One might prefer to maximize the expected minimum signal-to-noise ratio. If the three antennas have uncorrelated noises of equal powers one would then maximize $E(\psi_{\min})/(d + 2b)$. The dashed curve in Fig. 7 shows that $d = 0.45$, $b = 0.55$ for the maximizing detector.

APPENDIX C

The energy distributions (8) and (9)

It is convenient to have a special notation for real and imaginary parts of the field components at $P = 0$. S will always be a real part, J will always be an imaginary part. Subscripts 1, 2, or 3 on S or J denote the field, either $\varepsilon^{\frac{1}{2}} E_z$, $\mu^{\frac{1}{2}} H_x$, or $\mu^{\frac{1}{2}} H_y$. For example, the imaginary part of $\mu^{\frac{1}{2}} H_x$ is

$$J_2 = \sum_u u_y I(u).$$

Each of the six components S_1, J_1, S_2, J_2, S_3, J_3 is a linear combination of the $2N$ Gaussian variables $R(u_1), \cdots, I(u_N)$. Then these six variables have a joint Gaussian distribution, which is determined entirely by its 21-second moments.

All second moments of the form $E(S_i J_j)$ are zero. This follows because $E(R(u)U(v)) = E(R(u))E(I(v)) = 0$ for all N^2 choices of

u,v. Two other typical second moments are:

$$\begin{aligned} E(S_1S_3) &= E\{-\sum_{u,v} R(u)R(v)v_x\} \\ &= -\sum u_x E(R^2(u)) \\ &= -\sum u_x \\ &= -\sum \cos(2\pi k/N), \end{aligned}$$

and

$$\begin{aligned} E(S_2S_3) &= E\{-\sum_{u,v} R(u)R(v)u_y v_x\} \\ &= -\sum_u u_y u_x \\ &= -\tfrac{1}{2}\sum_{k=1}^{N} \sin(4\pi k/N). \end{aligned}$$

The identity

$$\sum_1^N \exp(ikt) = e^{it}\frac{1-\exp(iNt)}{1-\exp(it)}$$

can be used to prove that both are zero. In the first case set $t = 2\pi/N$ and take the real part (recall that $N \geqq 3$ is assumed). In the second case take $t = 4\pi/N$ and take the imaginary part. In like manner, one eventually finds that the only nonzero moments are

$$E(S_1^2) = E(J_1^2) = \sum_u 1 = N$$

$$E(S_2^2) = E(J_2^2) = \sum \sin^2\frac{2\pi k}{N} = \sum\left(\frac{1}{2} - \frac{1}{2}\cos\frac{4\pi k}{N}\right) = \tfrac{1}{2}N$$

$$E(S_3^2) = E(J_3^2) = \sum \cos^2\frac{2\pi k}{N} = \sum\left(\frac{1}{2} + \frac{1}{2}\cos\frac{4\pi k}{N}\right) = \tfrac{1}{2}N.$$

Note that these formulas hold only because $N \geqq 3$. The cases $N = 1$ and 2 are different, having $E(S_2^2) = E(J_2^2) = 0$ and $E(S_3^2) = E(J_3^2) = N$.

The six parts of the field components are independent Gaussian variables with joint probability density function

$$\tfrac{1}{2}(N\pi)^{-3}\exp(-\{S_1^2 + J_1^2 + 2(S_2^2 + J_2^2 + S_3^2 + J_3^2)\}/2N). \quad (23)$$

Now note $\psi(0) = dt_1 + bt_2$ where $t_1 = S_1^2 + J_1^2$ and $t_2 = S_2^2 + J_2^2 + S_3^2 + J_3^2$ are two independent variables with chi-squared distributions.

Then the desired distribution function is

$$F(\psi) = \text{Prob } (dt_1 + bt_2 \leqq \psi)$$
$$= \tfrac{1}{2}N^{-3}\int_0^{\psi/b} t_2 \exp\,(-t_2/N)\int_0^{(\psi-bt_2)/d} \exp\,(-t_1/2N)dt_1dt_2\,.$$

An elementary integration produces the final result (8).

The asymptotic form of $F(\psi)$ for small ψ can be obtained by differentiating (8) or, more simply, by the following argument. According to (23) the joint probability density of $S_1, J_1, S_2, J_2, S_3, J_3$ at the origin is $\frac{1}{2}(N\pi)^{-3}$. The inequality $\psi(0) \leqq \psi$ defines a small ellipsoid

$$d(S_1^2 + J_1^2) + b(S_2^2 + J_2^2 + S_3^2 + J_3^2) \leqq \psi$$

about the origin. This ellipsoid has six-dimensional volume

$$(\pi^3/6)d^{-1}b^{-2}\psi^3.$$

The probability that $(S_1, J_1, \cdots, J_3)$ lies in this ellipsoid is, apart from terms of higher order,

$$F(\psi) = \tfrac{1}{2}(N\pi)^{-3}(\pi^3/6)d^{-1}b^{-2}\psi^3$$
$$F(\psi) = (\tfrac{2}{3})d^{-1}b^{-2}(\psi/\bar{\psi})^3. \tag{24}$$

In Section VI, an additional wave, stronger than the others, was added to represent a "direct" wave. Let the direction u_0 of the direct wave be along the x-axis and let its amplitude be $A(u_0) = R$, a given real number. With $S_1, S_2, \cdots, J_3$ defined again to include the random fields only, the weighted energy density is

$$\psi(0) = d\{(R + S_1)^2 + J_1^2\} + b\{S_2^2 + J_2^2 + (R + S_3)^2 + J_3^2\}$$

with mean $\bar{\psi} = R^2 + 2N$. The asymptotic form of the distribution function for $\psi(0)$ may be derived in the same way as (24). Now the six-dimensional ellipsoid $\psi(0) \leqq \psi$ of volume $(\pi^3/6)d^{-1}b^{-2}\psi^3$ is centered on the point $(-R,0,0,0,-R,0)$. Equation (23) gives the probability density at that point and hence the result

$$F(\psi) = \psi^3/\{12N^3db^2 \exp\,(3R^2/2N)\} \tag{25}$$

approximately for small ψ.

APPENDIX D

Correlation coefficients

To prove (10) write,

$$E(\psi_1\psi_2) = 4dDE(\psi_E(0)\psi_E(P)) + 4dBE(\psi_E(0)\psi_H(P))$$
$$+ 4bDE(\psi_H(0)\psi_E(P)) + 4bBE(\psi_H(0)\psi_H(P)). \tag{26}$$

The four expectations on the right are:

$$E\{\psi_E(0)\psi_E(P)\} = N(N-1)\{1 + J_0^2(\beta r)\} + \tfrac{1}{4}NE(|A|^4) \quad (27)$$

$$\begin{aligned} E\{\psi_E(0)\psi_H(P) &= E\{\psi_H(0)\psi_E(P)\} \\ &= N(N-1)\{1 + J_1^2(\beta r)\} + \tfrac{1}{4}NE(|A|^4) \end{aligned} \quad (28)$$

$$\begin{aligned} E\{\psi_H(0)\psi_H(P)\} &= N(N-1)\{1 + \tfrac{1}{2}J_0^2(\beta r) + \tfrac{1}{2}J_2^2(\beta r)\} \\ &\quad + \tfrac{1}{4}NE(|A|^4). \end{aligned} \quad (29)$$

The proofs of (27), (28), (29) are alike. Only $E\{\psi_E(0)\psi_H(P)\}$ will be derived in detail.

Begin with

$$2\psi_E(0) = |\sum_u A(u)|^2$$

$$2\psi_H(P) = |\sum_u A(U)U \exp - i\beta U\cdot P|^2.$$

Then,

$$\psi_E(0)\psi_H(P) = \tfrac{1}{4}\sum_{u,v,U,V} A(u)A^*(v)U\cdot VA(U)A^*(V)\exp i\beta(V-U)\cdot P \quad (30)$$

with the summation variables u,v,U,V ranging over all N^4 ways of picking four vectors from $u_1, \cdots, u_N$. Now take the expectation of both sides of (30). Most of the N^4 terms in the sum have zero expectations because $E(A(u)) = E(A^2(u)) = 0$, and because $A(u)$, $A(v)$ are independent when $u \neq v$. Nonzero expectations can come from terms of three types:

(i) $N(N-1)$ terms with $v = u$, $V = U$, and $U \neq u$,
(ii) $N(N-1)$ terms with $V = u$, $v = U$, and $U \neq u$,
(iii) N terms with $u = v = U = V$.

The expectation of each term of type (i) is

$$\begin{aligned} E(|A(u)|^2U\cdot U\,|\,A(U)|^2) &= \tfrac{1}{4}E(|A(u)|^2)E(|A(U)|^2) \\ &= \tfrac{1}{4} \times 2 \times 2 = 1. \end{aligned}$$

All $N(N-1)$ terms contribute $N(N-1)$ to (28).

The expectation of each term of type (ii) is

$$\tfrac{1}{4}E\{|A(u)|^2|A(U)|^2u\cdot U \exp i\beta(u-U)\cdot P\}$$

$$= E\{u\cdot U \exp i\beta(u-U)\cdot P\}$$

Now let ϑ and Θ be the angles which u and U make with the vector P so that $u \cdot U = \cos(\vartheta - \Theta) = \cos\vartheta \cos\Theta + \sin\vartheta \sin\Theta$ and $(u - U)\cdot) = r\cos\vartheta - r\cos\Theta$. The expectation sought is

$$E\{\cos(\vartheta - \Theta) \exp i\beta r(\cos\vartheta - \cos\Theta)\}$$
$$= \left\{ \left| \int_0^{2\pi} \cos\vartheta \exp(i\beta r\cos\vartheta) d\vartheta/2\pi \right|^2 + \left| \int_0^{2\pi} \sin\vartheta \exp(i\beta r\cos\vartheta) d\vartheta/2\pi \right|^2 \right\}.$$

The two integrals may be recognized as Fourier coefficients in the well-known series

$$\exp(i\beta r\cos\vartheta) = J_0(\beta r) + 2\sum_{n=1}^{\infty} i^n J_n(\beta r)\cos n\vartheta.$$

Then each of $N(N-1)$ terms of type (ii) contributes $J_1^2(\beta r)$ to (28).

Each term of type (iii) is $\frac{1}{4}|A(u)|^4$; the total contribution to (28) of all N terms is $\frac{1}{4}NE(|A|^4)$.

APPENDIX E

Spectra

The normalized power spectrum of $\psi(V_0 t)$ is a Fourier transform of $E(\psi(0)\psi(V_0 t))/E(\psi^2(0))$. The expectations are obtainable from (10) with $D = d$, $B = b$, $P = V_0 t$ and from $E(\psi(0)) = E(\psi(V_0 t)) = 2N$. When N is large, one seeks the transform of

$$\frac{1 - d^2 J_0^2(\beta r) + 2\,db\,J_1^2(\beta r) + \frac{1}{2}b^2(J_0^2(\beta r) + J_2^2(\beta r))}{1 + d^2 + \frac{1}{2}b^2} \tag{31}$$

with $r = |V_0|\,t$.

The constant term in (31) represents the spectral line described in Section VIII. The remaining terms may be transformed using the equation

$$\int_{-\infty}^{\infty} J_n^2(x)\cos xy\,dx = \begin{cases} P_{n-\frac{1}{2}}(\frac{1}{2}y^2 - 1), & 0 < y < 2 \\ 0, & 2 < y < \infty \end{cases}$$

(Ref. 2, Erdelyi, *et al*, p. 46, transform 21), where $P_{n-\frac{1}{2}}(u)$ is the Legendre function of order $n - \frac{1}{2}$.

This result is applied in the form

$$\int_{-\infty}^{\infty} J_n^2(\beta \mid V_0 t \mid) \cos 2\pi f t \, dt = (-1)^n P_{n-\frac{1}{2}}(2\nu^2 - 1)/(\pi f_0)$$

for $0 \leqq \nu \leqq 1$ (recall (14) and $\nu = f/f_0$). One then obtains an expression for $s(f)$ which involves Legendre functions of orders $-\frac{1}{2}$, $\frac{1}{2}$, and $\frac{3}{2}$. This expression was not suitable for computing because there was no available table of Legendre functions of fractional order. However, the Legendre functions of half-integer order can be expressed as complete elliptic integrals by the following identities:

$$P_{-\frac{1}{2}}(x) = (2/\pi)K\{(\tfrac{1}{2} - \tfrac{1}{2}x)^{\frac{1}{2}}\}$$

$$P_{\frac{1}{2}}(x) = (2/\pi)\{2E((\tfrac{1}{2} - \tfrac{1}{2}x)^{\frac{1}{2}}) - K((\tfrac{1}{2} - \tfrac{1}{2}x)^{\frac{1}{2}})\}$$

$$(m+1)P_{m+1}(x) = (2m+1)xP_m(x) - mP_{m-1}(x).$$

(Ref. 1, Abramowitz and Stegun, Eqs. (8.13.1), (8.13.8), (8.13.11)). When the Legendre functions are replaced by elliptic integrals the Bessel function terms of (31) transform into

$$s(f) = \frac{2\{(3 - 4\nu^2 f^2)K((1-\nu^2)^{\frac{1}{2}}) + [(8\nu^2 - 4)b^2 - 12\,db]\,E((1-\nu^2)^{\frac{1}{2}})\}}{3\pi^2 f_0(1 + d^2 + \frac{1}{2}b^2)}. \quad (32)$$

The results (15) and (16) are special cases of (32).

REFERENCES

1. Abramowitz, M., and Stegun, I. A., (editors) *Handbook of Mathematical Functions*, Nat. Bu. Standards, Appl. Math. Series 55, 1964.
2. Erdelyi, A., Magnus, W., Oberhettinger, F., and Tricomi, F. G., *Tables of Integral Transforms, Bateman Manuscript Project*, I, McGraw-Hill, New York, 1954.
3. Ossanna, Jr., J. F., A Model for Mobile Radio Fading Due to Building Reflections: Theoretical and Experimental Fading Waveform Power Spectra, B.S.T.J., *43*, Nov., 1964, pp. 2935–2971.
4. Slack, M., The Probability Distributions of Sinusoidal Oscillations Combined in Random Phase, J.I.E.E., *93*, part III, 1946, pp. 76–86.
5. Young, Jr., W. R., Comparison of Mobile Radio Transmission at 150, 450, 900, and 3700 Mc, B.S.T.J., *31*, 6, 1952, pp. 1068–1085.

Contributors To This Issue

JOEL S. ENGEL, B.E.E., 1957, City College of New York; M.S.E.E., 1959, Massachusetts Institute of Technology; Ph.D., 1964, Polytechnic Institute of Brooklyn; Bell Telephone Laboratories, 1959–64; Bellcomm, 1964—. Mr. Engel first worked on the SAGE and BMEWS digital transmission systems. His later work at Bell Laboratories involved systems engineering studies of the toll telephone network. In October, 1964, he joined the Computer Systems Department at Bellcomm. Member, I.E.E.E. and Sigma Xi.

E. N. GILBERT, B. S., 1943, Queens College; Ph.D., 1948, Massachusetts Institute of Technology; Bell Telephone Laboratories, 1948—. He has worked on a variety of mathematical problems, most of which involve probability or combinatorial analysis. Member, American Math. Society.

DAVID L. JAGERMAN, B.E.E., 1949, Cooper Union; M.S., 1954, Ph.D., 1962, New York University; Bell Telephone Laboratories, 1964—. Mr. Jagerman has been engaged in mathematical research on stochastic command control systems, numerical quadrature theory, and mathematical properties of pseudo-random number generators. His recent work includes dynamic programming with application to optimal control systems.

T. T. KADOTA, B.S., 1953, Yokahama National University (Japan); M.S., 1956, and Ph.D., 1960, University of California (Berkeley); Bell Telephone Laboratories, 1960—. He has been engaged in the study of noise theory with application to optimum detection theory. Member, Sigma Xi and SIAM.

JACOB KATZENELSON, B.Sc., 1957, and M.Sc., 1959, Technion, Israel Institute of Technology; Sc.D., 1962, Massachusetts Institute of Technology; Bell Telephone Laboratories, 1962–1964. He is engaged in the analysis of nonlinear networks and simulation of electronic circuits on

digital computers. He is now with the Electronic Systems Laboratory and Project MAC at M.I.T. Member, IEEE, Tau Beta Pi and Sigma Xi.

KANEYUKI KUROKAWA, B.S., 1951, Doctor of Engineering, 1958, University of Tokyo; Assistant Professor, University of Tokyo, 1957–1963, Bell Telephone Laboratories, 1963—. He has been engaged in the design and development of microwave transistor amplifiers. Mr. Kurokawa presently supervises a group responsible for the development of microwave integrated circuits including transistor amplifiers. Member, IEEE, Institute of Electrical Communications Engineers of Japan and Institute of Electrical Engineers of Japan.

R. A. SEMPLAK, B.S., 1961, Monmouth College; Bell Telephone Laboratories, 1955—. He has been engaged in beyond-the-horizon radio propagation and three satellite communications projects: Project Echo, Telstar I and Telstar II. He has also participated in studies of the effects of rain on sky noise temperatures at 6-gc frequency and has recently completed an experimental study of the near-field Cassegrainian antenna. He is currently engaged in measuring the scattered radiation from various surfaces at 0.6-micron wavelength.

DAVID SLEPIAN, 1941–43, University of Michigan; M.A., 1947, Ph.D., 1949, Harvard University; Bell Telephone Laboratories, 1950—. He has been engaged in mathematical research in communication theory, switching theory, and theory of noise, as well as various aspects of applied mathematics. He has been mathematical consultant on a number of Bell Laboratories projects. During the academic year 1958–59, he was Visiting Mackay Professor of Electrical Engineering at the University of California at Berkeley. Member, AAAS, American Math. Society, Institute of Math. Statistics, IEEE, SIAM and U.R.S.I. Commission 6.

ESTELLE SONNENBLICK, B.A., 1933, Barnard College; M.A., 1934, Columbia University. Mrs. Sonnenblick came to Bell Laboratories as a programmer in 1960—. She has participated in the numerical solution of a great many Bell Laboratories problems. She began working with Mr. Slepian on the computation of prolate spheroidal wave functions in 1963. Member, Phi Beta Kappa.

UBERTO K. STAGG, B.S.E.E., 1958, Pennsylvania State University; M.S. in E.E., 1962, Ohio State University; Bell Telephone Laboratories, 1958—. Mr. Stagg has been engaged in No. 5 crossbar circuit develop-

ment and is presently involved in exploratory development of electronic adjuncts for No. 5 crossbar. Member, Tau Beta Pi and Eta Kappa Nu.

L. F. TRAVIS, B.S.E.E., 1958, University of New Hampshire; M.S.E.E., 1961, Northeastern University; Bell Telephone Laboratories, 1958—. At first engaged in the development of mobile radio main station equipment, Mr. Travis later worked on the development of carrier supply equipment for the L-type Multiplex, LMX-2. He currently supervises a group responsible for the equipment aspects of L Multiplex, T1 Carrier, and special wideband data terminals. Member, Tau Beta Pi and Phi Kappa Phi.

ROBERT E. YAEGER, B.S. in E.E., 1942, Worcester Polytechnic Institute; Bell Telephone Laboratories, 1942—. Mr. Yaeger's early work was concerned with carrier transmission systems and transistor circuit development. Subsequently, he was engaged in the early exploratory development and, later, the final development of the T1 carrier PCM system. Mr. Yaeger currently is Head, Data and Digital Systems Department in the Carrier Transmission Laboratory. Member, IEEE.

B.S.T.J. BRIEFS

An Observation Concerning the Application of the Contraction-Mapping Fixed-Point Theorem, and a Result Concerning the Norm-Boundedness of Solutions of Nonlinear Functional Equations

By I. W. SANDBERG

(Manuscript received July 20, 1965)

PART I

Let $\mathcal{B}$ denote a Banach space over the real or complex field $\mathcal{F}$. Let $\Theta(\mathcal{B})$ denote the set of (not necessarily linear) operators that map $\mathcal{B}$ into itself, with I the identity operator, and let $\| T \|$ denote the "Lipshitz norm" of T for all $T \in \Theta(\mathcal{B})$ (i.e.,

$$\| T \| \triangleq \sup_{\substack{x,y \in \mathcal{B} \\ \|x-y\| \neq 0}} \frac{\| Tx - Ty \|}{\| x - y \|}).$$

Observation:

Let A and B belong to $\Theta(\mathcal{B})$, and let $g \in \mathcal{B}$. Suppose that there exists $c \in \mathcal{F}$ such that (*i*) $(I + cA)^{-1}$ exists on $\mathcal{B}$, (ii) $\| A(I + cA)^{-1} \|$ and $\| B - cI \|$ are finite, and (iii) $\| A(I + cA)^{-1} \| \cdot \| B - cI \| < 1$. Then $\mathcal{B}$ contains exactly one element f such that $g = f + ABf$. (It can be verified that under our assumptions, $f \in \mathcal{B}$ satisfies $g = f + ABf$ if and only if f satisfies

$$g = f + A(I + cA)^{-1}[(B - cI)f + cg].)$$

For the special case in which A is a linear operator, this result is well known* and has been applied often in the engineering literature [see, for example, Ref. 2]. The fact that it can be generalized as indicated suggests that the scope of its range of applicability to engineering problems can be extended significantly.

* The linearity of A plays an essential role in all of the previous proofs known to this writer. See, for example, Ref. 1.

PART II

Let $\mathcal{K}$ denote an abstract linear space, over the real or complex field $\mathfrak{F}$, that contains a normed linear space $\mathcal{L}$ with norm $\| \cdot \|$. Let Ω denote a set of real numbers, and let P_y denote a linear mapping of $\mathcal{K}$ into $\mathcal{L}$ for each $y \in \Omega$, such that $\| P_y h \| \leqq \| h \|$ for all $h \in \mathcal{L}$ and all $y \in \Omega$. We say that a (not necessarily linear) operator T is an element of the set Θ if and only if T maps $\mathcal{K}$ into itself and $P_y T = P_y T P_y$ on $\mathcal{K}$ for all $y \in \Omega$. The symbol I denotes the identity operator on $\mathcal{K}$.

Proposition:†

Let A belong to Θ, and assume that A maps the zero-element of $\mathcal{L}$ into itself. Let B map $\mathcal{K}$ into itself. Let $f \in \mathcal{K}$, and let $g = f + ABf$. Suppose that there exists $\lambda \in \mathfrak{F}$ such that

(i) $(I + \lambda A)$ is invertible on $\mathcal{K}$, $(I + \lambda A)^{-1} \in \Theta$, and $A(I + \lambda A)^{-1}$ maps $\mathcal{L}$ into itself

(ii) $\eta_\lambda \triangleq \sup \{ \| A(I + \lambda A)^{-1} h \| / \| h \| : h \in \mathcal{L}, h \neq 0 \} < \infty$

(iii) there exists a nonnegative constant k_λ and a function $p_\lambda(y)$ with the property that

$$\| P_y(B - \lambda I) f \| \leqq k_\lambda \| P_y f \| + p_\lambda(y) \text{ for all } y \in \Omega$$

(iv) $\eta_\lambda k_\lambda < 1$.

Then

$$\| P_y f \| \leqq (1 - \eta_\lambda k_\lambda)^{-1} [(1 + |\lambda| \eta_\lambda) \| P_y g \| + \eta_\lambda p_\lambda(y)]$$

for all $y \in \Omega$.

Proof:

Let $y \in \Omega$. Then, since $Bf = (I + \lambda A)^{-1}[(B - \lambda I) f + \lambda g]$, we have

$$\begin{aligned} P_y f &= P_y g - P_y A (I + \lambda A)^{-1} [(B - \lambda I) f + \lambda g] \\ &= P_y g - P_y A (I + \lambda A)^{-1} P_y [(B - \lambda I) f + \lambda g], \end{aligned}$$

and hence

$$\begin{aligned} \| P_y f \| &\leqq \| P_y g \| + \eta_\lambda \| P_y [(B - \lambda I) f + \lambda g] \| \\ &\leqq \| P_y g \| + \eta_\lambda \| P_y (B - \lambda I) f \| + |\lambda| \eta_\lambda \| P_y g \| \\ &\leqq (1 + |\lambda| \eta_\lambda) \| P_y g \| + \eta_\lambda k_\lambda \| P_y f \| + \eta_\lambda p_\lambda(y), \end{aligned}$$

which establishes the proposition.

† This proposition is a generalization of a result proved in Ref. 3, and is of considerable utility in stability studies of nonlinear physical systems.

Comments:

Consider the important special case in which: $\mathcal{K}$ denotes the set of real-valued locally-square-integrable functions on $[0, \infty)$, $\mathcal{L}$ denotes the space of real-valued square-integrable functions x on $[0, \infty)$ with norm

$$\| x \| = \left(\int_0^\infty x(t)^2 dt \right)^{\frac{1}{2}},$$

$\Omega = [0, \infty)$, and P_y is defined by

$$\begin{aligned} (P_y h)(t) &= h(t), \qquad t \in [0, y] \\ &= 0, \qquad t > y \end{aligned}$$

for all $h \in \mathcal{K}$. Suppose that A is defined on $\mathcal{K}$ by

$$(Ah)(t) = k_0 h(t) + \int_0^t [k_1(t - \tau) + k_2(t - \tau)] h(\tau) d\tau$$

for all $h \in \mathcal{K}$, where k_0 is a real constant, k_1 and k_2 are real-valued measurable functions on $[0, \infty)$, with k_1 bounded on $[0, \infty)$ and k_2 integrable on $[0, \infty)$.

Let

$$K(s) = k_0 + \int_0^\infty [k_1(t) + k_2(t)] e^{-st} \, dt$$

for $\sigma \triangleq \text{Re}[s] > 0$, and, with λ a real constant, assume that

$$\sup_{\sigma > 0} \left| \frac{K(s)}{1 + \lambda K(s)} \right| < \infty.$$

Then, with the aid of some known results[4] from the theory of Fourier transforms, it can be proved that

(i) $(I + \lambda A)^{-1} \in \Theta$, and $A(I + \lambda A)^{-1}$ maps $\mathcal{L}$ into itself,
(ii) there exists a zero-measure subset $\mathfrak{N}$ of $[0, \infty)$ such that

$$\lim_{\sigma \to 0+} \frac{K(\sigma + i\omega)}{1 + \lambda K(\sigma + i\omega)}$$

exists for all $\omega \in \tilde{\mathfrak{N}} \triangleq [0, \infty) - \mathfrak{N}$,

and

$$\text{(ii)} \quad \eta_\lambda \triangleq \| A(I + \lambda A)^{-1} \| = \operatorname*{ess\,sup}_{\omega \in \tilde{\mathfrak{N}}} \left| \lim_{\sigma \to 0+} \frac{K(\sigma + i\omega)}{1 + \lambda K(\sigma + i\omega)} \right|.$$

These facts can be used to extend some of the results of Ref. 3 to a more

general class of integral equations. For example, let B denote the mapping of $\mathcal{K}$ into itself defined by the condition that $(Bh)(t) = b(t)h(t)$ for all $t \geqq 0$ and all $h \varepsilon \mathcal{K}$, where $b(\cdot)$ is a real-valued measurable function with the property that there exist real numbers α and β such that $\alpha \leqq b(t) \leqq \beta$ for all $t \geqq 0$. With $g \varepsilon \mathcal{L}$, let $g = f + ABf$ with $f \varepsilon \mathcal{K}$. Let k_1 be a constant, and let

$$K(s) = k_0 + s^{-1}k_1 + \int_0^\infty k_2(t)e^{-st}dt$$

for all $s \varepsilon \mathcal{S} \triangleq \{s: s \neq 0, \sigma \geqq 0\}$. Suppose that

$$\begin{aligned} &1 + \tfrac{1}{2}(\alpha + \beta)k_0 \neq 0 \\ &1 + \tfrac{1}{2}(\alpha + \beta)K(s) \neq 0 \qquad \text{for all } s \varepsilon \mathcal{S}, \end{aligned} \tag{1}$$

and

$$\tfrac{1}{2}(\beta - \alpha) \sup_{\omega > 0} \left| \frac{K(i\omega)}{1 + \tfrac{1}{2}(\alpha + \beta)K(i\omega)} \right| < 1. \tag{2}$$

Then, an application of the proposition shows that $f \varepsilon \mathcal{L}$. This result, which is concerned with feedback loops containing a pure integrator, cannot be proved as an application of the result similar to our propositions given in Ref. 3, because there A is assumed to map $\mathcal{L}$ into itself.

REFERENCES

1. Anselone, P. M., *Nonlinear Integral Equations*, Univ. of Wisconsin Press, 1964, pp. 152–153.
2. Sandberg, I. W., On Truncation Techniques in the Approximate Analysis of Periodically Time-Varying Nonlinear Networks, IEEE Trans. Ckt. Theory, *CT-11*, June, 1964, p. 195.
3. Sandberg, I. W., Some Results on the Theory of Physical Systems Governed by Nonlinear Functional Equations, B.S.T.J. *44*, May–June 1965, p. 871.
4. Titchmarsh, E. C., *Introduction to the Theory of Fourier Integrals*, Oxford University Press, 1948, pp. 125, 128.