VOLUME XXX JANUARY 1951 NO. 1

THE BELL SYSTEM TECHNICAL JOURNAL

DEVOTED TO THE SCIENTIFIC AND ENGINEERING ASPECTS OF ELECTRICAL COMMUNICATION

50¢ per copy

$1.50 per Year

THE BELL SYSTEM TECHNICAL JOURNAL

Published quarterly by the
American Telephone and Telegraph Company
195 Broadway, New York 7, N. Y.

Leroy A. Wilson, *President* — Carroll O. Bickelhaupt, *Secretary* — Donald R. Belcher, *Treasurer*

SUBSCRIPTIONS

Subscriptions are accepted at $1.50 per year. Single copies are 50 cents each. The foreign postage is 35 cents per year or 9 cents per copy.

PRINTED IN U. S. A.

The Bell System Technical Journal

Vol. XXX | *January, 1951* | *No. 1*

The Type N-1 Carrier Telephone System: Objectives and Transmission Features

By R. S. CARUTHERS

(*Manuscript Received Oct. 17, 1950*)

The N1 Carrier System is a 12-channel, double-sideband system for single cable application. It provides low loss, stable, high velocity service for toll and exchange circuits in the range from 15 or 20 miles to 200 miles. Units and subassemblies are miniaturized and arranged on a plug-in basis. Emphasis has been placed on reduction in cost of components, as well as simplification of manufacturing methods, engineering, installation and maintenance. Economy is achieved by many novel features, principal among which is a built-in low cost compandor. By compressing and expanding the volume range of speech, the compandor permits much higher tolerance of noise and crosstalk, thereby substantially lowering the cost of both line and terminal facilities. Other important features are self-contained dialing and supervisory signaling, an individual channel regulator, and automatic equalization through the use of "frequency frogging," or interchange of high- and low-frequency groups at each repeater.

Introduction and General Technical Description

THE N-1 Carrier System is the most recent addition to the alphabetic list of carrier telephone systems which began in 1918 with the A system. This and many other systems produced since then have passed into obscurity. Others like the C, H, J, K, and L Systems* carry the majority of telephone traffic for distances exceeding 100 miles. Even though carrier has been the backbone of all the longer-haul telephone service in the country, these systems, and in particular the terminals, have been too expensive for short-haul use. This has prevented tapping the great mass of circuits owned largely by the Associated Companies and extending into nearly every city and town. The M system, developed primarily for power line use, has found limited application in this field.

The objective in the design of the N system has been to provide a single cable carrier facility which, without special cable treatment, will be economical for distances as short as 15 to 20 miles and which will be technically satisfactory in performance for a nominal maximum of 200 miles. Relaxa-

* See list of references at end of article.

tion of requirements, made possible by limiting the system to 200 miles (instead of the usual transcontinental 4000 miles), has been an important factor in helping to meet the low-cost objective. Elimination of the need for two cables permits use of a large part of the five million miles of toll and exchange circuits in Associated Company voice frequency cables for carrier operation.

An important aim of type N carrier is directed at the exchange plant where, even though the mileage is less, the number of circuits for exceeds that in toll use. Here the benefits of a high grade carrier facility are numerous. Exchange Plant makes use of a large percentage of small-gauge, high-capacitance cables, heavily loaded to reduce the net loss. Economically it is difficult to apply carrier or voice repeaters to these relatively short-length circuits. Many circuits to suburban points are now routed over toll trunks, because of the high loss of the exchange type circuits. Type N can be used to afford a low-loss, high-grade exchange circuit which can be switched in the manner usual for tandem and similar circuits. Low-loss trunks employing type N will be of great benefit in the ever increasing distances in the suburban areas between satellite points and their outlets. Direct, high-grade trunk groups, always at a premium and first selection of automatic switching equipment, can be increased in number.

Another important objective, in addition to the provision of a stable, low-loss, high-velocity talking circuit, is that of providing built-in signaling arrangements suitable for dialing and supervision. Such a system is urgently needed to meet the rapidly expanding demands of toll line dialing, as well as for exchange circuits. Such a signaling channel has been made available at a frequency just above and directly associated with the voice channel which it serves.

The emphasis placed on economy in the development of the N system has not resulted in a marked lowering of standards of performance. On the contrary, the N system, within its range of operation, sets new standards in many respects, notable among which is stability of net loss. The objectives have been met rather by a combination of new approaches and new circuit elements, backed up by closely coordinated cooperative effort in cost reduction of components, assemblies, and finally, of the complete system.

Among the many features making possible such a development as N1 Carrier, certain are outstanding. The most important of these is the compandor, a device for compressing and expanding the volume range of speech, thereby affording an improvement in the amount of noise and cross-talk which can be tolerated. The effects of the compandor are far reaching, both in the line and in the terminals. The need for expensive line treatment, such as crosstalk balancing, is eliminated; band filter discrimination can be reduced; and signal levels can be raised without undue interference.

The N system employs a cable pair in each direction. In order to operate in a single cable the two directions are further separated by the use of different frequency bands; 44–140 kc for one direction on one pair; and 164–260 kc for the other direction on the other pair. Double-sideband, carrier transmitted operation, very similar to that of a radio system, is used, with channels spaced 8 kc apart. The voice channel bandwidth is 250–3100 cycles. The dialing and supervisory control frequency is at 3700 cycles.

Frequency frogging, involving interchange and inversion of frequency bands at each repeater, is accomplished by modulation with a 304 kc carrier, and serves two important purposes: Circulating crosstalk paths around the repeater are blocked; and the system is made self-equalizing for as many as ten repeater sections, having a gross loss of between 400 and 500 db.

Either paired or quadded, 16, 19, 22, or 24-gauge cable conductors can be employed, with suitable variation in repeater spacing. The nominal spacing of repeaters is 8 miles for 19-gauge and 6 miles for 22-gauge conductors. No limitation is placed on the percentage of cable conductors on which N carrier can be applied in a toll cable. Accordingly, as many as 1800 channels can be obtained from a 300-pair cable. For built up connections, two N systems can operate in tandem to make up a toll trunk. At the most, not more than 4 to 6 links of N are expected in tandem in a long multilink connection.

Many additional transmission features are listed and briefly described in Table I.

Frequency Allocation

The frequency allocation of the system is shown in Fig. 1. In order to coordinate system frequencies in the same cable some with odd numbers of repeaters, some with even numbers, and some circuits starting or stopping at intermediate repeater points of other systems, it is necessary to arrange the terminals to transmit and receive either high or low group frequency bands. The channel modulators and demodulators in the terminals, however, use carriers only in the high group band at 8 kc intervals between 168 and 256 kc. Thus, when transmitting high group frequencies to the line and receiving low group frequencies, the high group transmitting unit (HGT) merely amplifies the channel frequencies. The associated low group receiving unit (LGR) however, employs a group modulator with 304 kc carrier that inverts the received low group of line frequencies into the upper band for channel separation in the receiving channel band filters. When transmitting low group to the line and receiving high group, the group modulator and 304 kc oscillator are used in the transmitting side of the circuit. Similarly, in the repeater, low and high group bands are interchanged between input and output lines through use of group modulators with 304 kc carriers.

Choice of the one group alone for primary modulation and demodulation of the speech bands stems largely from the desire to use only 12 designs of

TABLE I

TRANSMISSION FEATURES OF N1 CARRIER TELEPHONE SYSTEM

1. Built-in compandor affording an effective signal-to-noise improvement of 20–25 db.
2. Frequency frogging and inversion to improve crosstalk and furnish automatic equalization.
3. Built-in signaling equipment in each channel to provide supervision and dial pulsing. Tone on-tone off operation employing 3700 cycles.
4. Message channel bandwidth 250–3100 cycles. Transmission of special services (telegraph and telephoto) through standard message channel equipment. 3500 cycle program channel plus 11 message channels, or 5000 cycle program channel plus 9 message channels provided by special program channel equipment.
5. Automatic regulation of each channel at the receiving terminal by the individual channel carrier.
6. All alarms built in with special carrier system failure alarm operating on transmitted carriers automatically freeing subscriber dial equipment.
7. Use of noise generator where needed to mask intelligible crosstalk and obtain satisfactory performance in exchange type cables.
8. Built-in resistance hybrid arrangements for 2-wire termination at non-gain switching points or alternative use of 4-wire termination at -16 and $+7$ levels for standard interconnection to existing broadband intertoll carrier systems. As much as $+10$ level is permissible for special purposes.
9. Repeaters spaced at 8-mile intervals on 19-gauge toll cable and at shorter distances on high-capacity or smaller-gauge exchange cable.
10. Power fed to pole mounted repeaters 8 miles (19-gauge toll cable) on either side of an office repeater, thus requiring power supply stations about 24 miles apart. Power is fed over the cable pairs by simplex connection and use of $+130$ volt and -130 volt batteries.
11. Automatic regulation of line repeaters by thermistor flat gain adjustment, controlled by total output power of the 12 transmitted carriers.
12. In service switching of repeater and terminal circuits.
13. Small, lightweight, portable transmission measuring equipment for office and pole cabinet use.
14. Simple order wire and alarm equipment provided to alarm power failures at unattended power offices and to permit communication with all repeater points.

channel band filters rather than 24. Easier filter requirements, occasioned by the use of double-sideband operation and the compandor, together with the fact that, for the high group, all harmonics fall outside the useful band, result in the elimination of the need for transmitting band filters. Thus, the only filter needed in the 12-channel group of sidebands and carriers is a common filter to suppress transmission of speech sidebands on harmonics of the channel carriers. An important factor in the choice of the high group for receiving channel band filters was the better performance obtained in the simple radio type slug-tuned coils in this frequency range, and the smaller size of condensers needed for tuning.

Compandor

While the compandor principle is not new, it is believed that, for the first time, full advantage of the compandor has been taken in the design of a carrier system. To assist in explaining these advantages, general compandor principles will be reviewed in the light of the present development. The 1A compandor[1], designed more than ten years ago, has had considerable usage in open-wire carrier systems in reducing crosstalk, but in the N

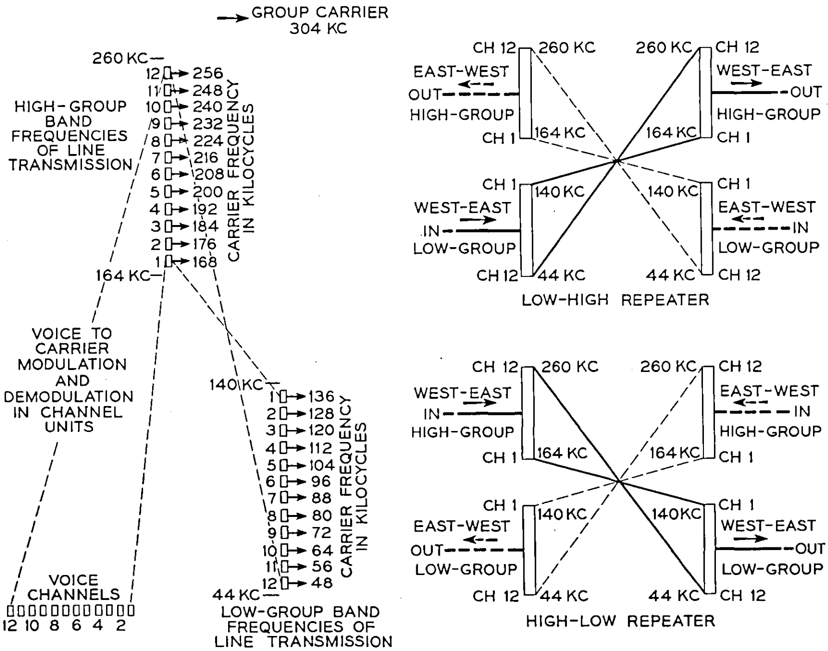


Fig. 1—N-1 carrier frequency allocations for terminals and repeaters.

system a more compact and cheaper unit was needed with requirements revised to match the reduced maximum length of circuits. The word compandor is a contraction of compressor and expandor—the compressor in the transmitting terminal compressing the input range of speech volumes for passage over a wire or radio transmission medium where a variety of noise and crosstalk interferences are present—the expandor in the receiving terminal expanding the received range of compressed speech volumes to the original range. A 20–28 db noise advantage is derived, and can be explained as follows: Weak speech volumes most susceptible to system disturbances are lifted and carried at higher level over an intervening noisy medium.

[1] Application of Compandors to Message Circuits, C. W. Carter, Jr., A. C. Dickieson and D. Mitchell—*A.I.E.E. Trans.*, Vol. 65, pp. 1079–1086.

The stronger volumes need less increase in proportion to the volume. When the circuit is idle 28 db gain is introduced by the compressor and 28 db loss by the expandor. Any disturbance in the transmission medium in the absence of speech receives 28 db attenuation in the expandor. Loss is removed from the expandor as the speech volume increases and the noise increases correspondingly. In a well designed compandor with proper time constants the increased noise will tend to be continuously masked by the increased speech volume. Interferences to the listener during silent speech periods, such as intelligible crosstalk or audible tones, receive the full 28 db of noise suppression in the expandor. Interference in the presence of speech receives less than full suppression in the expandor the stronger the speech. Table II shows test results of noise advantage of the N-1 compandor at several noise values and speech volumes.

In Fig. 2(a) a level diagram shows the gain and loss introduced by the compressor and expandor for signals of different strengths. A signal of 5 db

TABLE II
COMPANDOR NOISE ADVANTAGE

Thermal Noise (dba at 0 level)	Speech Volume at 0 Level			
	None	−30VU	−10VU	0 VU
53	28.0	24.7	24.0	20.3
58	27.0	22.2	22.2	19.9
63	24.0	20.0	17.8	17.2
68	18.5	17.8	14.6	13.7

above 1 milliwatt (+5 dbm) is shown as unmodified by compressor and expandor. A signal input to the compressor of −50 dbm receives 27.5 db gain and the resulting −22.5 dbm signal input to the expandor receives 27.5 db loss. For each signal input to the compressor weaker than +5 dbm by 2 db, the compressor introduces 1 db more of gain and the expandor 1 db more of loss to a maximum of 28 db gain and loss respectively at −51 dbm input to the compressor or −23 dbm input to the expandor. In Fig. 2(b) input vs output is plotted for compressor, expandor and the combination. The slopes of these input-output plots are 1/2 for compressor and 2/1 for expandor.

In Fig. 3, (a) and (b), compressor and expandor circuit schematics are shown for the N-1 Carrier System. Compressor input and expandor output are connected to the resistance hybrid circuit at the left of the compressor schematic for conversion to the 2-wire voice circuit input. Alternative connections for 4-wire operation and an adjustable gain control to establish over-all circuit net loss also are shown. Input voice signals to the compressor pass through the germanium variolosser, are compressed to half the input

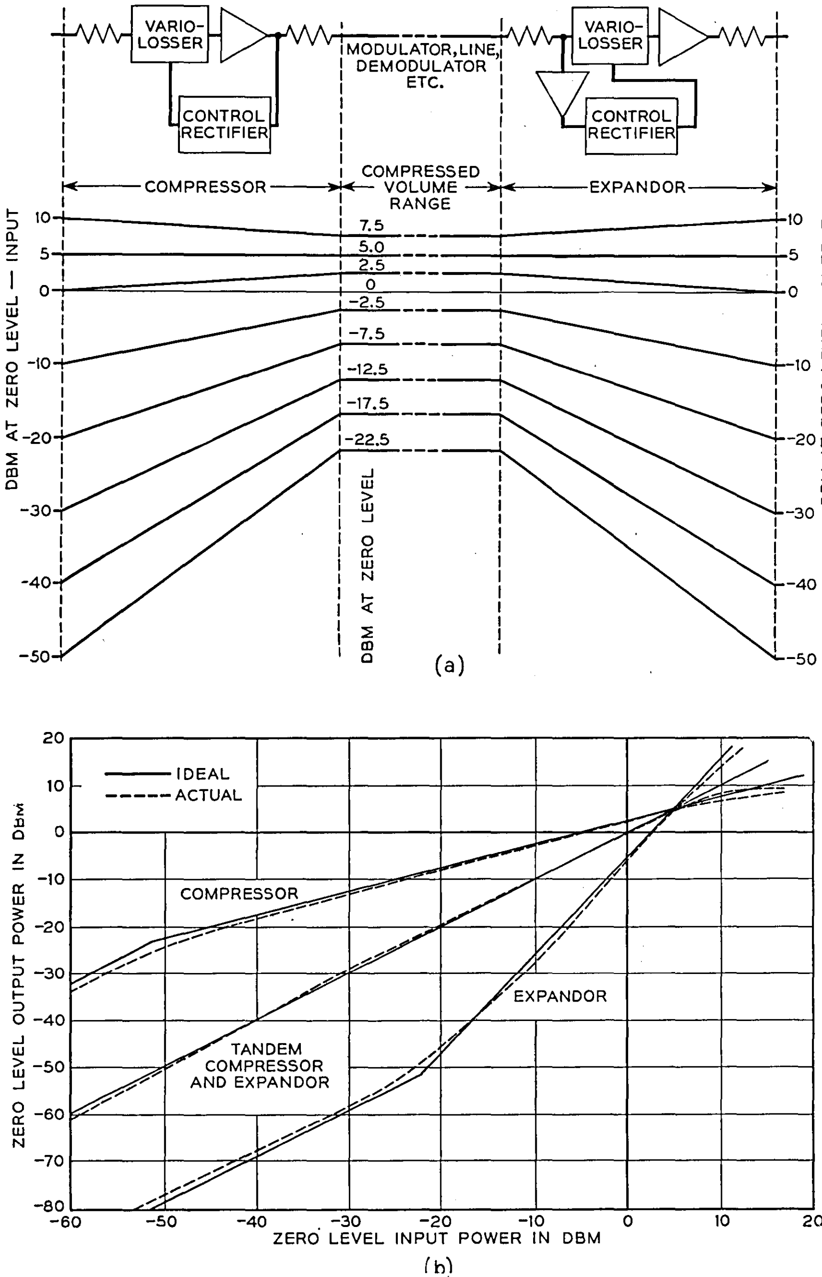


FIG. 2 (a)—Compandor action on steady tones of different levels.
(b)—Input-output load characteristics of N-1 compandor.

volume range, then are amplified in a 2-stage feedback amplifier. Most of the amplifier output is fed through the control circuit where it is rectified in a germanium bridge circuit. A condenser-resistance filter in the control

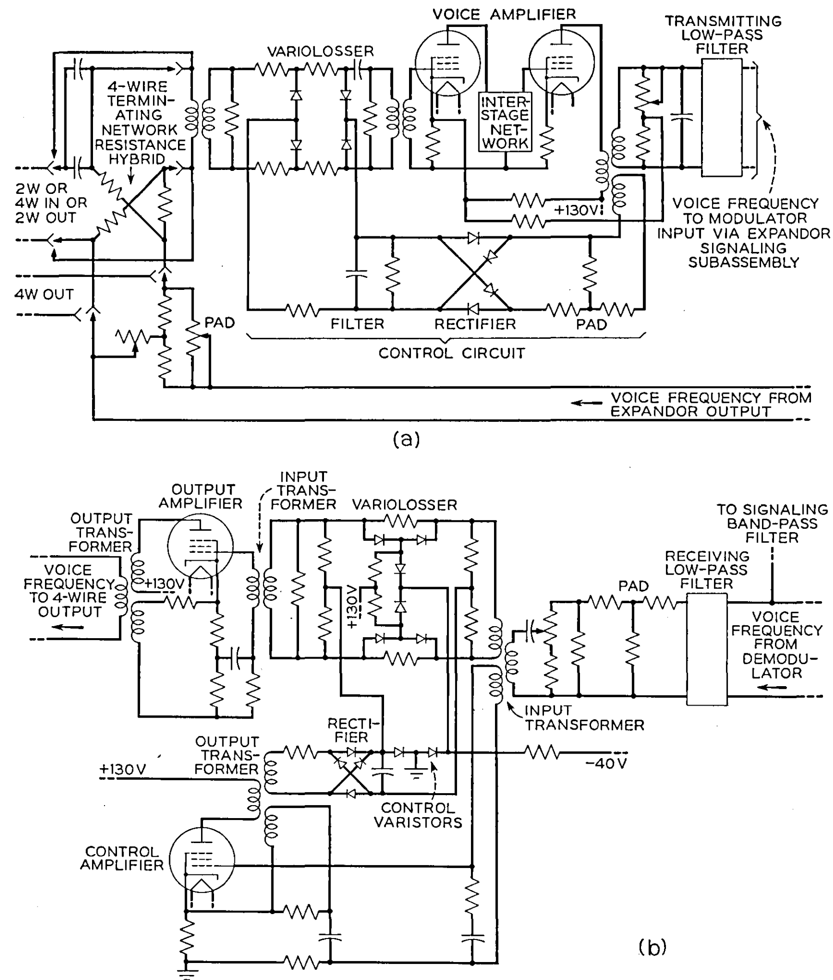


Fig. 3 (a)—Channel compressor circuit.
(b)—Channel expandor circuit.

circuit output passes only the rectified syllabic envelope of the speech frequencies. The control circuit filter output current is fed to the midpoints of the germanium variolosser bridge arrangement. The time constants of the control circuit filter are chosen for fast attack (3–5 milliseconds)[2]

[2] 90% of final value reached in time indicated.

to prevent syllabic speech bursts from overloading circuits following the compressor, and for slow recovery (30–50 milliseconds)[2] so that the variolosser will introduce fixed loss over a syllabic interval. Too slow a recovery time also is harmful, tending to leave the expandor at low loss after the speech burst is over. Thus, the noise can be heard at the end of each syllable.

The expandor in Fig. 3(b), like the compressor, consists of a variolosser, an output amplifier, and a control circuit which rectifies the compressed range of speech signals. Thus, the expandor control circuit is operated by the expandor input speech signal and is called forward acting, while the compressor control circuit is operated by the compressor output signals and is called backward acting. The rectified syllabic envelope control circuit currents in the compressor and expandor are made as near alike as possible through choice of like circuit constants and levels, so that good tracking of compressor and expandor variolossers will result.

Integration of the compandor into the design of the N-1 system from the start has yielded many advantages both from a line standpoint, and in repeater and terminal circuit and equipment design. A listing of these advantages follows:

Line

Operation to frequencies as high as 260 kc without need for far-end crosstalk balancing. Crosstalk in cable increases about 6 db as the frequency doubles and the ability to balance crosstalk becomes rapidly unsatisfactory above 60 kc. The N system with satisfactory crosstalk for 200 miles would be satisfactory for only one or two miles without compandors.

Repeater spacing can be about 25 db longer (40% more miles) than with no compandor, without limitations from near-end crosstalk or line noise.

Less precise balance in line and equipment against longitudinal noise can be tolerated.

Longitudinal noise suppression coils are eliminated in voice pairs not used for carrier at repeaters in telephone offices.

Reflected near-end crosstalk requirements are eased markedly, thus permitting much less precise equipment impedances.

Repeater

Poorer modulation can be tolerated, thus allowing 25 db less feedback, 25 db less non-regenerative gain and fewer repeater tubes. As many as 25 repeaters can be tolerated in tandem. Without the compandor even one repeater would make the system unsatisfactory from this standpoint.

Repeater directional filter discrimination requirements are reduced by about 25 db.

Use of small and cheap filter coils and repeater transformers with permalloy cores is possible without harmful modulation.
Less precise transformer impedance and balance requirements, in conjunction with reduced size, eliminates the need of electrostatic shields between windings.

Terminal

Aids in elimination of transmitting channel band filter, and in large reduction of receiving band filter requirements.
Permits higher levels of carrier, speech and signaling tone without intolerable noise, crosstalk or interchannel modulation effects.

Equipment

Much more freedom is allowed in equipment layout and wiring, permitting more compacting, miniaturizing and less use of shield plates, shielded cans, and shielded wiring, without harmful noise pickup and crosstalk couplings.
Operation is feasible from common office battery with large reduction in individual circuit filtering. Signaling and speech circuits can be used on the same office battery without need for separate office wiring, fusing and alarms.

Frequency Frogging

Like the compandor, frequency frogging is vital to the N system, and numerous benefits result from its use. Primarily the purpose was to eliminate interaction crosstalk, i.e., crosstalk from the output of one repeater into a paralleling voice pair and thence back into the input of other repeaters. In K carrier cables this crosstalk path was eliminated by using two cables and at a repeater point connecting one cable to repeater inputs and the second cable to repeater outputs. The voice pair passing by the repeater point and remaining in the one cable thus was not exposed to both repeater inputs and outputs. In the N system in a single cable a modulator in each repeater frogs the frequency band from low group to high group, and in the following repeater back again from high group to low group. Thus, repeater outputs are always in one frequency band, and repeater inputs in the other, so that the crosstalk through the paralleling voice path can always be blocked by a filter at the repeater input. This approach is invaluable in N carrier where the alternative to frequency frogging is to use a second cable or to add suppression filters in all the paralleling voice pairs. In Fig. 4, cable frogging in K carrier and frequency frogging in N carrier are illustrated diagrammatically.

In addition to frequency frogging, the two frequency bands are inverted in passing through the N repeater. Thus, the highest frequency channel in one line section becomes the lowest frequency channel in the succeeding line section. So nearly constant are the sums of the losses in two line sections for all channels for the frequency range chosen, that equalization is provided without resort to any major slope correction in the repeaters. The small amount of slope and bulge remaining are easily taken care of in the repeater through use of a few shaping elements in the feedback circuit.

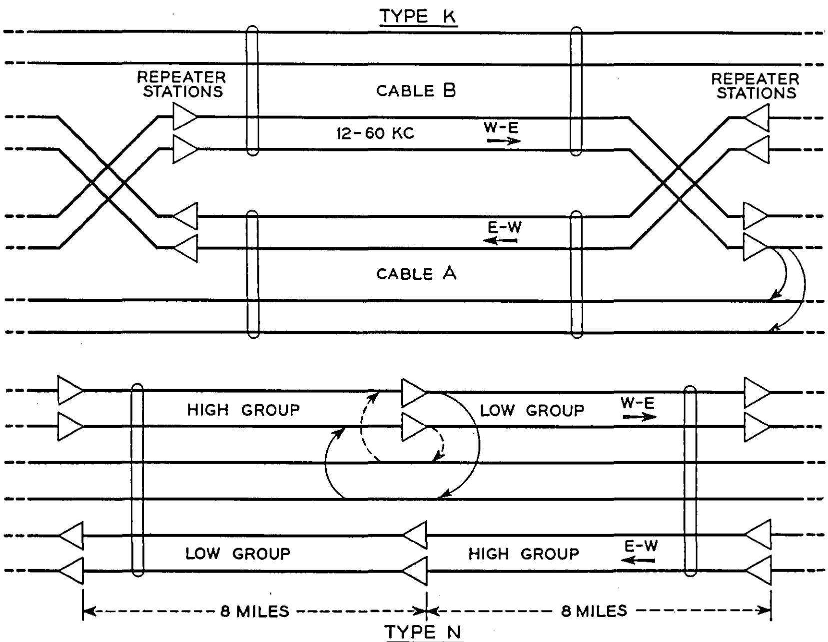


FIG. 4—Cable frogging and frequency frogging.

In Fig. 5(a) the sum of the line losses is shown for two successive 7.5 mile cable sections. The residual slope of only 3.7 db is in contrast to about 34 db of slope in two successive low-group sections in an unfrogged system. The flat line loss of about 90 db is accompanied by only about 0.4 db of bulge. Also shown in Fig. 5(a) is the summation of LH and HL repeater gains. The difference in slope between line and repeater amounts to about 1.5 db and is taken care of by a small range slope control in the repeater.

The remaining difference between line and repeater is nearly flat with frequency and is compensated for either through use of flat pads in the line (span pads) or through use of the repeater flat gain regulators. At about each tenth repeater enough frequency distortion has accumulated through lack of match between repeater and line to require use of a deviation equalizer. The anticipated shape of this characteristic, shown in Fig. 5(b), is based on 19-gauge cable and the deviations among the first 46 factory made

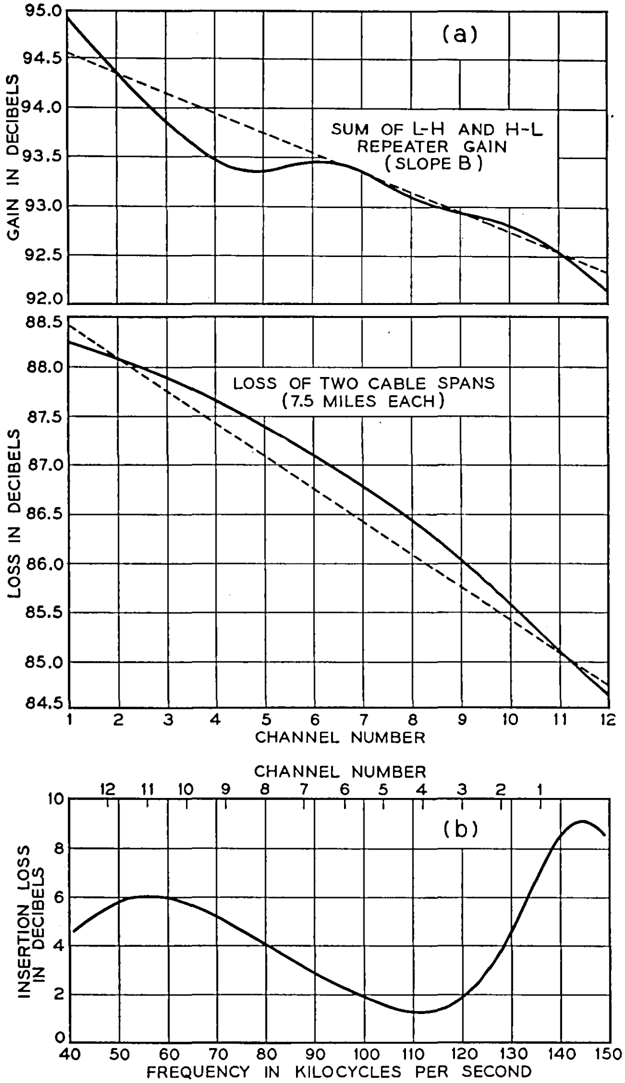


FIG. 5(a)—Repeater gain and cable loss characteristic.
(b)—Loss-frequency characteristic of deviation equalizer.

repeaters. More accurate determinations can be made by measurement of early systems in the field.

Because of frogging the maximum repeater gain is that required at the mean frequency instead of maximum line frequency. Furthermore, average repeater spacings required in opposite directions of transmission are identical. In a system without frequency frogging repeaters and using different frequency bands in the two directions, the average repeater spacings required in the two directions would be quite different. In the latter system, for various controlling reasons repeaters are installed at the closer spacing of the high-frequency group for both directions of transmission.

Frequency frogging and inversion also benefit regulation. With temperature change, a first order change in flat line loss occurs, a second order change occurs in slope and a considerably smaller bulge change occurs. In so far as the two successive line sections are at the same temperature, slope changes are nearly compensated. The bulge changes are so slight that they impose but small regulation requirements in the channel equipment even in a long system.

Signaling

Each voice frequency channel of the N system contains in addition to the compressor and expandor circuit, a signaling circuit operating at 3700 cycles for both dialing and supervision. Physically both the signal sending and receiving circuits are included in the expandor plug-in subassembly. The compressor and the carrier subassemblies, together with the expandor-signaling subassembly, comprise the channel terminal unit. Compressor and expandor-signaling subassemblies are alike and interchangeable among the 12 N channels. Carrier subassemblies differ only in respect to the oscillator and filters.

Signaling over the N system may assume a variety of forms. In ringdown operation 1000-cycle ringing signals may pass over the voice channel without need for the N signaling circuit, or, on the other hand, ringing may be converted to d-c. and passed over the N system by turning its 3700 cycle signal tone on and off. In dial operation digits are carried as in the national toll-line dialing plan either by multifrequency key pulsing or by dial make and break connections of a single-frequency tone. With multifrequency key pulsing the two frequency combinations of 700, 900, 1100, 1300, 1500 and 1700 cycles pass directly through the voice channel. Supervisory on-hook and off-hook signals, as well as dial pulse signals, are carried by turning 3700 cycles on and off. In addition, called numbers may be transmitted over the N system either by the operator verbally or by recorded methods of the panel call announcer system. Revertive pulsing and panel call indicator signals are not provided for in the N system, there being little need for N systems in panel offices where this equipment is used.

The N carrier signaling system uses 3700 cycle tone just out of the voice band. The carrier channel bandwith is made wide enough to carry both the speech band and signal tone. Use of 1600-cycle tone of the national toll dialing plan would impair the noise advantage of the compandor unless arrangements were used to by-pass the signal tone around the compandor. Because of the complexity and cost of these arrangements, 1600-cycle signaling is used over the N system only in special cases. The N signal tone is injected after the compressor in the transmitting terminal and removed before the expandor at the receiving end of the system. Thus, the compandor is left free to operate on speech signals of various levels without interference and consequent impairment of its noise advantage by the signal tone. Use of the compandor permits an unusually high level of signal tone (0 dbm at 0 level in contrast to -20 dbm at 0 level for 1600-cycle signaling) without interference into the message circuits. In the compressor circuit [Fig. 3(a)] the compressor low-pass filter cuts the speech band off at about 3100 cycles, preventing speech interference to the signal channel. In the expandor circuit [Fig. 3(b)] the expandor low-pass filter passing the voice band blocks the signal tone from the expandor.

In Fig. 6, (a), (b) and (c), the three parts of the signal circuit are shown—the signal oscillator circuit, the keyer circuit and the signaling receiver. The 3700-cycle oscillator is a resistance condenser Wien bridge type using a thermistor for stabilization of the output. One oscillator serves for the signaling supply of 12 channels. It is housed in the low-group subassembly of the group terminal unit. A low-impedance output circuit makes it possible to key on one channel or remove channels without disturbing others.

In Fig. 6(b) the signal keyer circuit is shown. On-hook or off-hook signals are received over the M lead from the trunk as ground and battery, respectively. With ground on the M lead the bias on germanium varistors in the keyer bridge becomes positive, which connects the 3700-cycle oscillator to the channel modulator through the keyer transformer. With -48 volt battery off-hook signal on the M lead, the varistors receive negative bias and 3700-cycle transmission to the modulator is blocked.

The signaling receiver [Fig. 6(c)] is connected to the output of the carrier channel demodulator in multiple with the expandor low-pass filter input [Fig. 3(b)]. It consists of a receiving band filter about 150 cycles wide, an amplifier stage, a limiter, a cathode follower impedance converting stage, a germanium varistor rectifier, a delay circuit, a d-c. amplifier and an output relay. The band filter is made as narrow as practicable to reject noise and still pass sidebands of the dial pulses without perceptible distortion, with allowances for shift in the signal oscillator frequency and manufacturing variation of the band filter. The multivibrator limiter gives constant amplitude square wave 3700-cycle output for any 3700-cycle signal input to the receiver up to about 7 db below normal level and well above normal level.

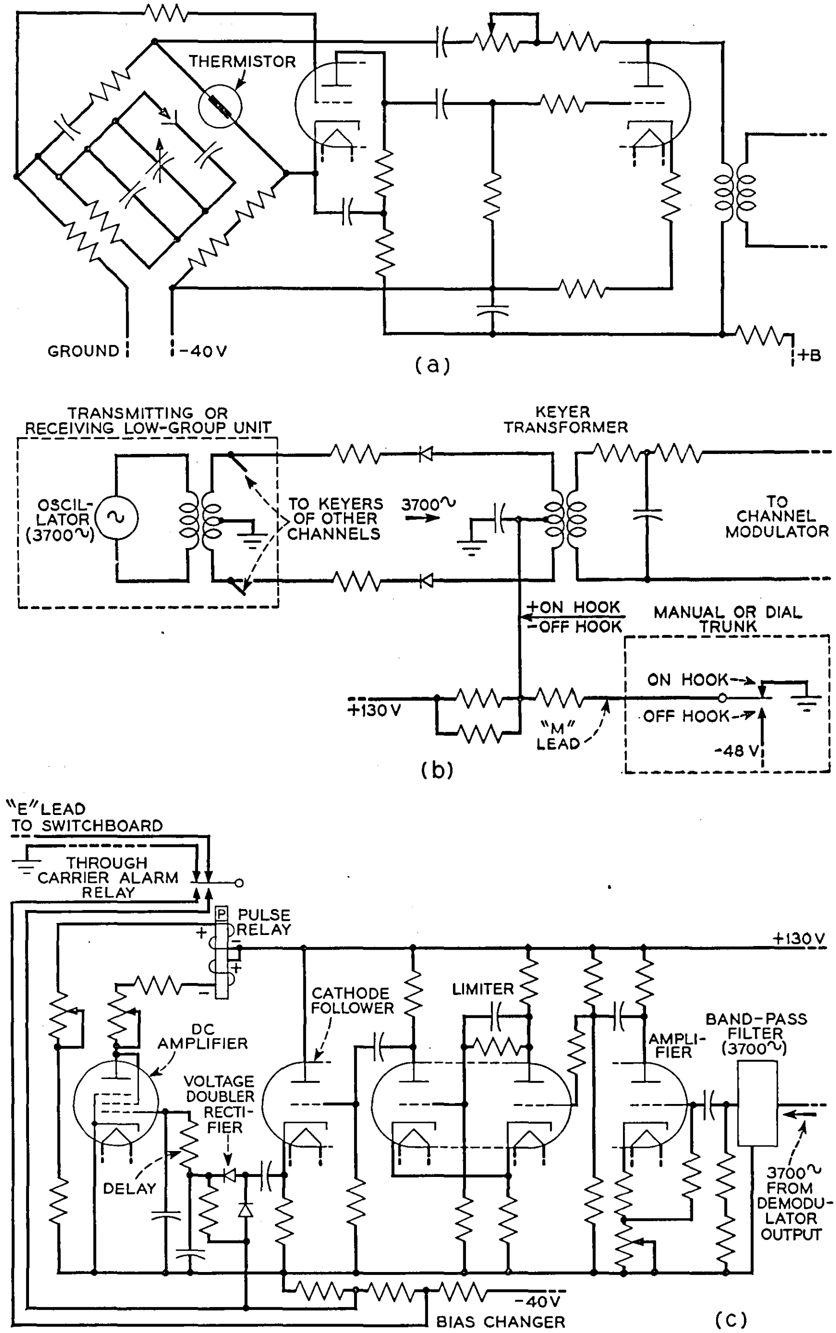


FIG. 6(a)—3700 cycle signaling tone oscillator circuit.
(b)—Channel signaling keyer circuit.
(c)—Channel signaling receiver circuit.

Such amplitude stabilization is needed to prevent excessive pulse distortion from the slow rise and fall of pulse edges in the delay circuit. Only during the transition periods between make and break can noises more than 7 db below normal signal level cause pulse distortion. The cathode follower supplies low impedance drive to the germanium rectifier which, with its low resistance load, results in highly stable operation. The delay circuit functions primarily to discriminate against short duration noise bursts during the off-hook signal condition. Unwanted relay operation at this time would flash the operator or cause registration of a wrong number. The output relay, which is of the mercury contact type, has a split winding for ease of adjusting the per cent break and to minimize pulse distortion with 130-volt battery variation. Bias change on the varistor rectifiers through the relay contacts prevents excessive first pulse distortion despite the high amount of delay used in the delay circuit. Control potentiometers in the circuit adjust "just operate" sensitivity, relay current and per cent break. On-hook and off-hook are sent from the signal receiver to the trunk circuit as open and ground, respectively, on the "E" lead.

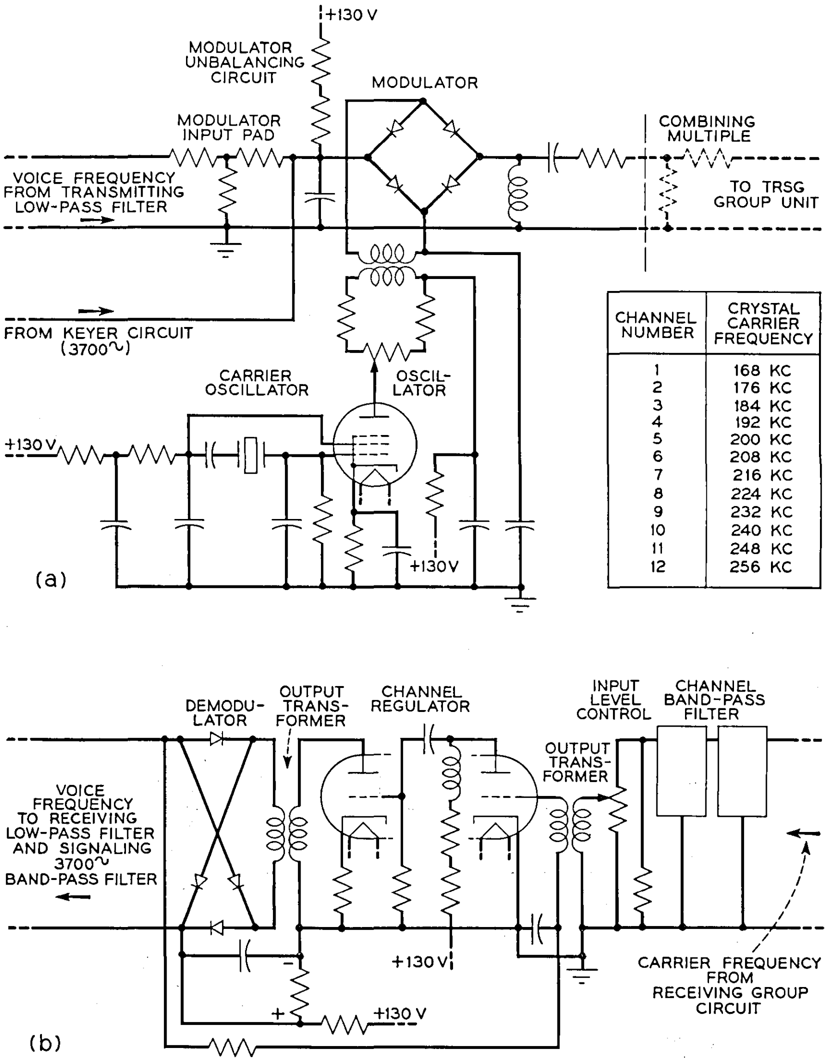
Use of the N system as a part of the nationwide networks of dialing intertoll trunks imposes strict limits on the amount of pulse distortion contributed. The severest requirements are when operating into step-by-step offices where pulses distorted beyond per cent break limits of about 44–72 per cent may produce wrong numbers. A typical connection from a dial and outgoing trunk over a two-link N connection into a step-by-step office might be expected to have a distribution of pulse distortion as follows:

	Normal Battery and Normal System Levels (% Break)
Dial	63.5 ± 4%
Outgoing Trunk	−6 ± 2%
1st Type N Link	+1 ± 2%
Pulse Link	−1 ± 1.5%
2nd Type N Link	+1 ± 2%
	58.5 ± 11.5%

During extreme battery and level conditions on one N circuit link about ± 2% more pulse distortion can be expected.

Carrier Frequency Transmitting and Receiving Circuit

The third part of the channel unit is the carrier subassembly. It contains a germanium varistor modulator and individual channel crystal carrier oscillator in the transmitting circuit and the channel band filter, an automatic gain control or channel regulator and a germanium varistor demodulator in the receiving circuit. Figs. 7(a) and 7(b) show the carrier channel circuits. At the input of the transmitting circuit, the output of the compressor LPF and the signal keyer are connected. At the output all



CHANNEL NUMBER	CRYSTAL CARRIER FREQUENCY
1	168 KC
2	176 KC
3	184 KC
4	192 KC
5	200 KC
6	208 KC
7	216 KC
8	224 KC
9	232 KC
10	240 KC
11	248 KC
12	256 KC

FIG. 7(a)—Channel carrier frequency transmitting circuit.
(b)—Channel carrier frequency receiving circuit.

modulators are connected together through pads individual to each channel and mounted on the terminal frame. The crystal oscillator uses a pentode tube arranged as a triode oscillator with electron coupling to the plate for output. Because of the large spread between input and output signal bands the balanced series modulator arrangement can use a simple condenser input and inductor output filter. Poor and uncontrolled carrier leak would result in change in amplitude and phase of the transmitted carrier used to regulate each channel in the receiving circuit and to demodulate the two sent sidebands of voice and signal frequencies. Selection of germanium units in manufacture provides good balance of the varistor bridge and results in very low level of carrier leak. A controlled amount of d-c. is introduced at the modulator input terminals from the 130-volt battery and sends carrier at adequate level and correct phase for demodulation of the speech sidebands in the demodulator without over modulation on the speech peaks. The transmitted carrier amplitude is stable for channel regulation purposes, varying in amplitude only with the 130-volt battery which, in most offices, will seldom be outside of one per cent limits.

In the receiving carrier channel the band filter at the input selects the particular channel from the receiving group unit output. The crystals in the oscillator and the band filters are the only differences among the 12 channels for this subassembly. The filter design is such that one or more channel units can be removed from service without affecting the performance of the others.

At the output of each band filter a potentiometer is used to adjust for the transmission inequalities from one channel to the next of the over-all system. The two-stage double-triode regulator automatically adjusts for any subsequent changes in level of the received channel signals. The regulator is operated by d-c. obtained from carrier rectification in the linear demodulator. Typical delayed AVC action is obtained by biasing out part of the rectified carrier voltage with a part of the 130-volt battery supply. The time constant of the regulator circuit is slowed to about 5 seconds to provide adequate regulating speed without false operation on speech or line hits. The demodulator is operated as a linear detector with highest peaks on speech sidebands just below 100% modulation. The unusually high impedance level of 10,000 ohms for the varistor demodulator provides large d-c. voltage for the gain control circuit without undue instability. This results from inclusion of the demodulator in the mu circuit of the regulator. The output of the demodulator is connected to the signaling band filter and the expandor LPF.

Terminal Transmitting and Receiving Group Units

The group units serve essentially as terminal repeaters for transmitting and receiving directions. There are four varieties: low-group transmitting (LGT) and associated high-group receiving units (HGR) or high-group transmitting (HGT) and associated low-group receiving units (LGR). A terminal uses three plug-in subassemblies: the low group unit, the high group unit, and the oscillator unit containing 304 kc and 3700 cycle signaling oscillators. The oscillator unit is always associated with the low-group unit.

In Fig. 8, (a), (b) and (c), typical group unit circuits are shown. The 12 channels combine in the resistance pads of the multiple on the terminal frame and enter the transmitting group unit through the E filter. This filter, as previously described, is used in both types of transmitting units to suppress transmission on harmonics of the channel carrier frequencies. The noise generator at the input terminals supplies tube noise of equal amplitude at each high-group channel frequency to mask intelligible crosstalk in short-toll or exchange circuits where the system noise may be low and the crosstalk disturbing. A potentiometer controls the noise magnitude, which is always set well below the tolerable limit. A slope equalizer in the high band is used in the transmitting group unit to produce a compromise slope among the channels at repeater points. Through its use input and output levels of low-high and high-low repeaters are sloped either positively or negatively by about 7 db, so that no one channel has more than 7 db disadvantage from a noise and modulation standpoint. A compensating slope equalizer in the receiving group unit restores the channels to flat band at its output. The group modulator and 304 kc oscillator are alike in group units and repeaters, whether used for low- to high-group band translation or vice versa. The crystal oscillator differs only in minor respects from the channel unit oscillator. The modulator is a double balanced ring type using four $\frac{3}{16}$ inch diameter copper oxide discs. Group and repeater modulators employ copper oxide to minimize noise. Low signal levels and high carrier level are used in the group modulators to produce low interchannel modulation. As a result good balance in the modulator and an output filter are needed to suppress carrier leak. The two-stage feedback amplifier is alike in repeaters and group units except in minor respects.

Repeater

A block schematic of a repeater station is shown in Fig. 9. The plug-in repeater is shown within the dashed vertical lines. Mounted on the frame and permanently wired are span pads, artificial line sections and deviation equalizers when needed. The deviation equalizer is used only at low-group frequencies and is placed in repeater input or output, whichever results in locating the equalizer in an office instead of a pole cabinet. Resistance span pads in 2 db steps from 2 db to 24 db are used to build out line sections shorter than 8 miles at channel 1 frequency to 46 db loss in the low group and 50 db in the high group. For line sections shorter than four miles, artificial cable sections in two-mile and four-mile sizes are available. In

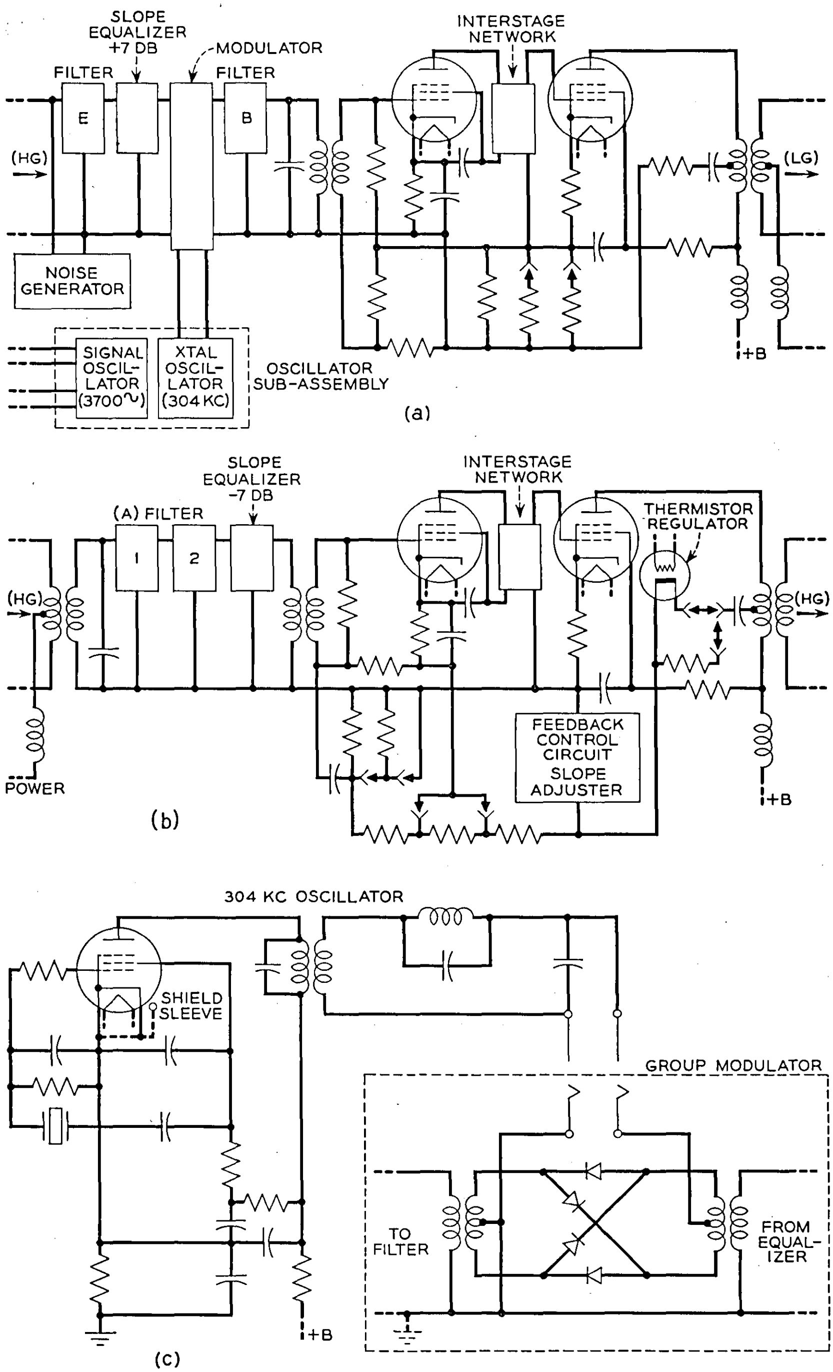


Fig. 8(a)—Terminal low group transmitter circuit.
(b)—Terminal high group transmitter circuit.
(c)—Group modulator and 304 KC oscillator circuit

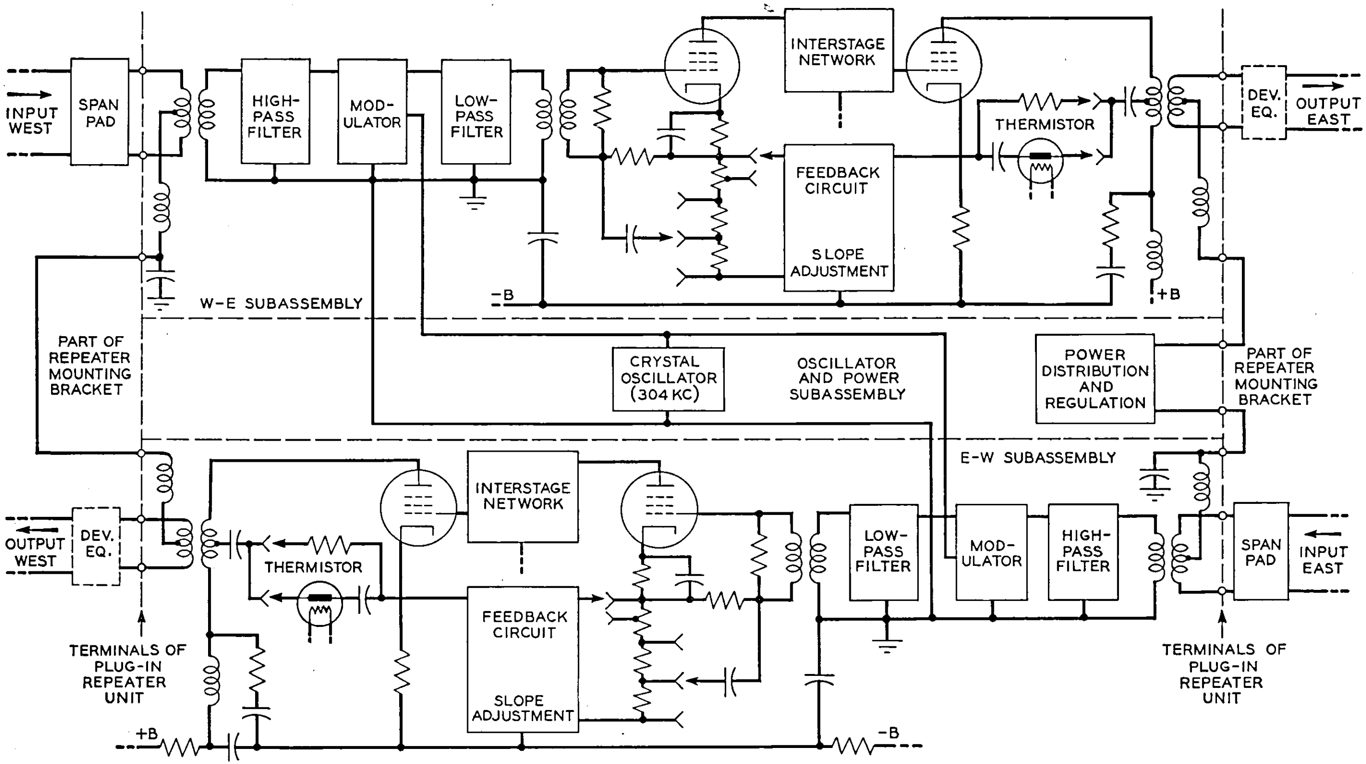


Fig. 9—N-1 Repeater block schematic.

most cases very little or no span pad would be used and only in rare cases would the artificial line sections be needed. The small amount of slope needed to supplement the span pad loss is provided in a three-step repeater slope control. A total of about seven db of slope range within about 1 db accuracy is available in two successive repeaters. This would permit compensation for about three miles of inequality in successive line sections. The slope changes are produced by switching simple networks in the amplifier feedback circuit.

The plug-in repeater consists of three separate subassemblies connected together through the multicontact plug and jack connectors used in all N equipment units. Of these the east-west and west-east modulator-amplifier units are alike electrically and mirror images mechanically. The oscillator-power subassembly forms the third subassembly. These units are alike in LH and HL repeaters. The amplifier-modulator circuits supply the transmission paths through the repeater and differ between LH and HL repeaters only in whether the input circuit accepts low-band frequencies and amplifies high-group frequencies at the output or the reverse. The input A or C filter blocks near-end crosstalk from line frequencies within the same quad, which would overload and produce interchannel interference within the same system, or it blocks interaction crosstalk through tertiary-voice pairs from other systems in the cable. The output B or D filter suppresses 304 kc carrier leak and the unwanted upper sideband on the 304 kc carrier of the input group frequency band.

In addition to the slope control two additional controls are used in the feedback circuit of the repeater amplifier. Resistance strapping options made in the factor adjust each manufactured repeater to a nominal gain of 48 db to an accuracy of 1 db. A second control of flat gain is obtained from a thermistor directly heated by a fraction of the total repeater output power. Inasmuch as the signal tone power, when present in a channel, is about 12 db below the channel carrier power, and the speech of an average talker about 15 db below the power of the carrier, the total output power is almost entirely carrier. As the line changes in attenuation with temperature, the change in the strength of the carriers at the repeater output supplies more or less heat to the thermistor pellet. An increase in heat due to decreased line loss at cold temperatures causes the thermistor resistance to decrease and produce more amplifier feedback and less repeater gain, thus offsetting the change in line loss. The normal operating range of the thermistor resistance is from about 1000 to 20,000 ohms. At nominal gain of 48 db the resistance will always be about 9000 ohms. Each pellet is controlled in manufacture to have its nominal resistance value for a specified amount of repeater output power. The temperature at which the pellet operates for the standard repeater output power would vary appreciably except for provision of an ambient temperature compensating circuit. This circuit consists of a heater electrically insulated from the pellet and controlled by a disc thermistor at the repeater temperature. As the repeater temperatures decreases, the disc thermistor increases its resistance, allowing more current to flow into the heater winding. The thermistor pellet is

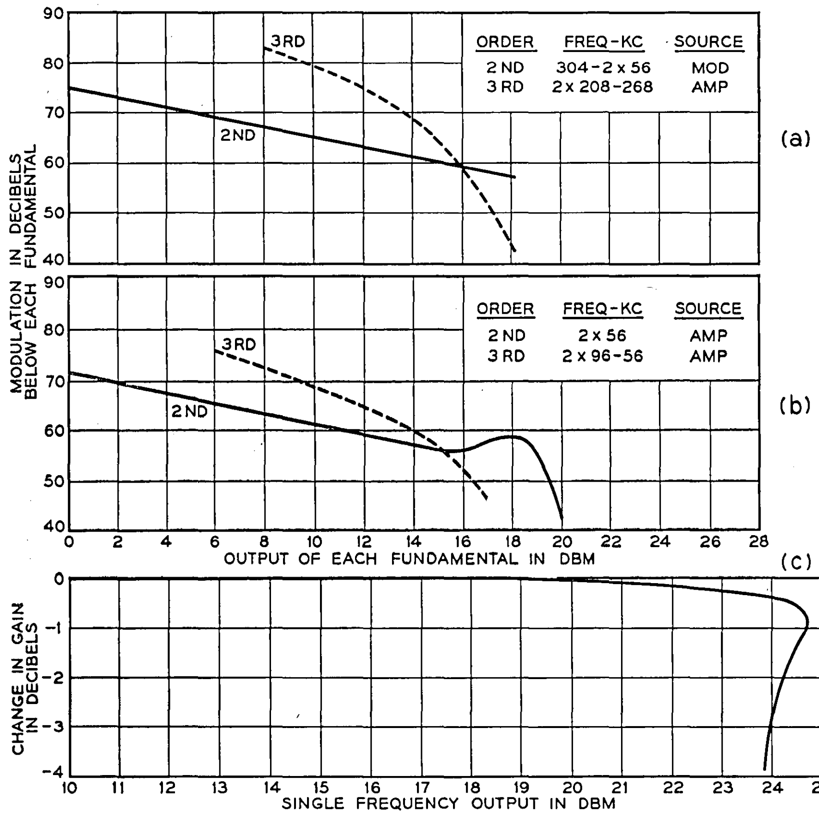


FIG. 10(a)—Low-high repeater modulation.
(b)—High-low repeater modulation.
(c)—Gain-load characteristics of high-low and low-high repeaters.

maintained at a nominal or thermostated temperature which, in general, is from about 135° to 185°F. This allows operation of the regulator with repeater temperatures from about −20°F to 130°F with little change in its control range. Beyond these temperatures the performance deteriorates slowly.

The low level operation of the modulator and repeater amplifier combined with the high level of carrier in the modulator and large amplifier feedback, result in low interchannel modulation. In Fig. 10, (a) and (b), one- and two-frequency modulation curves are shown. Figure 10(c) shows the single-frequency load characteristic. Most of the modulation crosstalk results from the many third-order combinations of carriers and speech sidebands in a repeater. Third-order products of this type add in phase in a string of repeaters. Twenty repeaters are 26 db worse in modulation crosstalk than one repeater.

Repeater and Terminal Levels

The operating levels of the system are all referred to the strengths of the individual carriers which are each made 15 db above one milliwatt (+15 dbm) at reference 0 level of one sideband. Thus, a + 5 dbm signal at 0 level

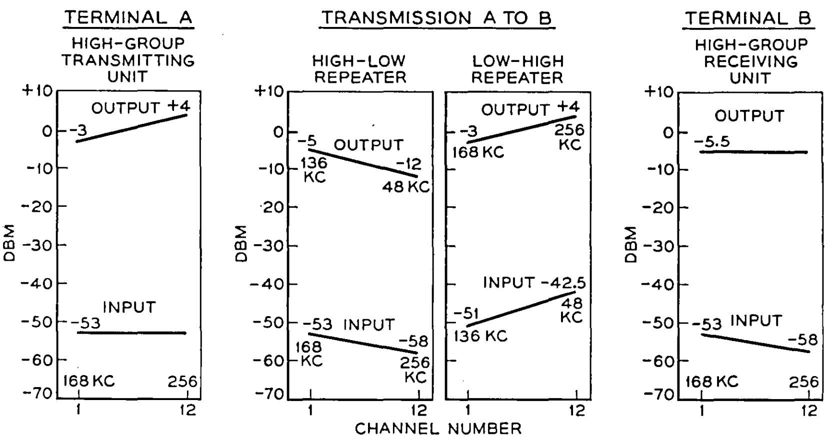


Fig. 11(a)—N-1 repeater and group unit level diagrams.

in the voice circuit is modulated to produce two sidebands each of +5 dbm at 0 single sideband level in the carrier part of the system.

Six db of loss must be inserted between 0 single sideband level and 0 voice level at the output because of the in-phase addition of the two sidebands upon demodulation. Thus, the two +5 dbm sidebands become +5 dbm at 0 voice level in the output. Zero dbm of 3700 cycle tone is used for signaling at 0 level and, since it is inserted after the compressor and removed before the expandor, each sideband is 15 db below the carrier. Limiting of speech peaks in the compressor restricts maximum values to about +9 dbm at 0 level. In-phase addition of the maximum speech sideband peaks nearly 100% modulate the carrier (+ 15 dbm at 0 level).

In Fig. 11, (a) and (b), carrier level information is given for repeaters and terminals. High-group output repeaters and terminals have carrier outputs sloped from −3 to +4 dbm with a total of +12 dbm of carrier

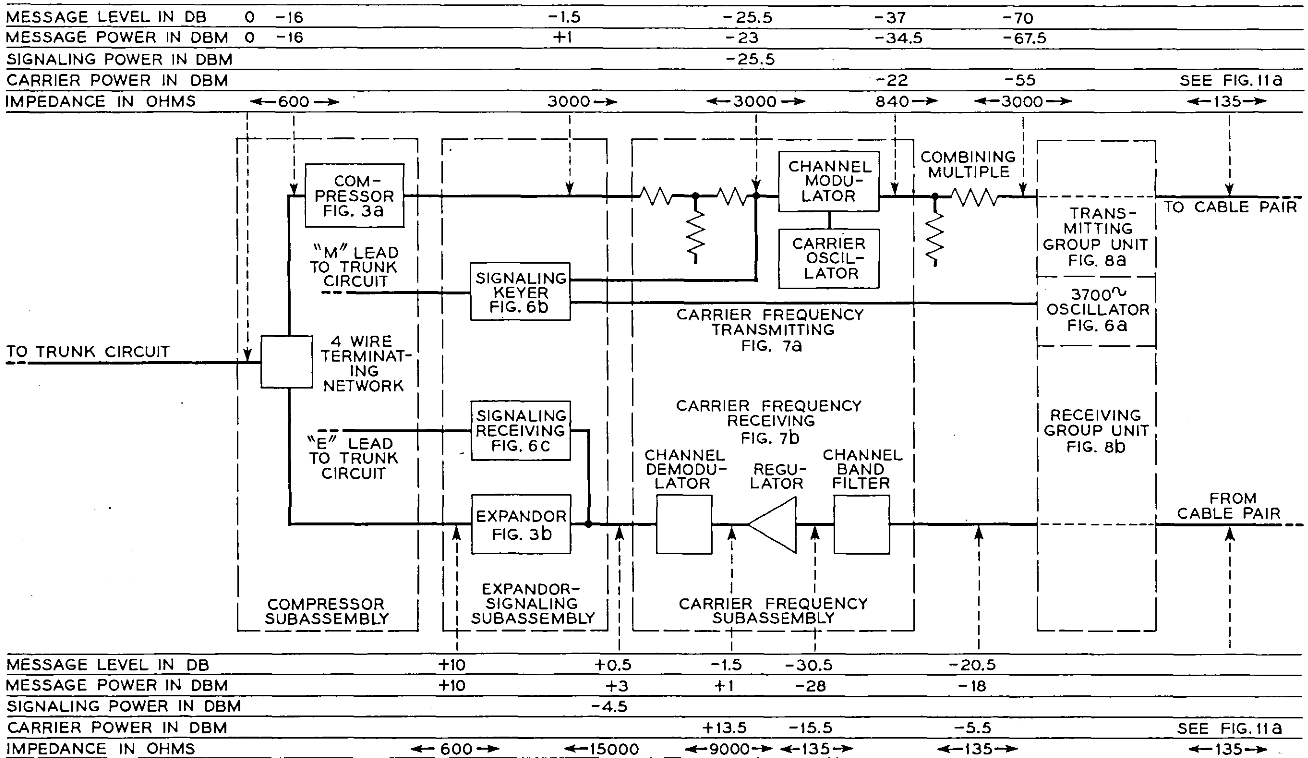


FIG. 11(b)—N-1 channel unit simplified schematic and level diagram.

power. Low-group output repeaters and terminals have carrier outputs sloped from −12 to −5 dbm with total carrier power of +3 dbm. Transmitting group input carrier levels of −53 dbm and receiving group output carrier levels of −5.5 dbm are alike in low-and high-group terminal units.

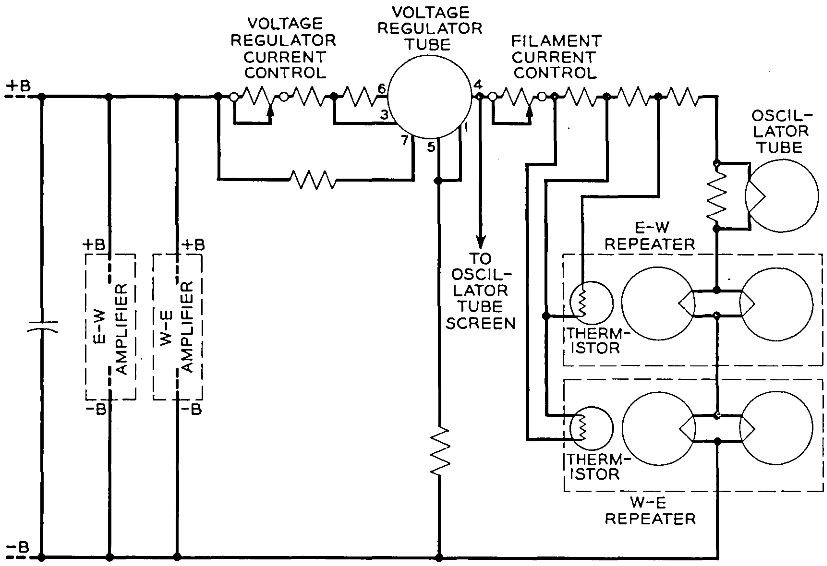


FIG. 12—Repeater power regulator—Distributer Circuit.

Strengths of signals at various levels before and after the compressor can be simply calculated from the following equation:

$$P_C = \frac{P - L}{2} + L_C + 2.5,$$

where P and L are the power in dbm and level in db before the compressor and P_C and L_C the corresponding values after the compressor.

POWER SUPPLY

The power supply at terminals is obtained from +130 volt and −48 volt batteries. All vacuum tubes have 20-volt heaters used two in series across an adjusted 40 volts from the 48-volt battery. Repeaters are supplied either locally from +130 volt battery or over the simplexes of the cable quad to pole-mounted points using +130 volt and −130 volt battery. A series parallel arrangement of tube and thermistor heaters in the repeater obtain closely regulated voltage from a gas tube control circuit. This arrangement is shown in Fig. 12(a). So closely are the tube heater voltage and current held despite changes in supply voltage, cable conductor resistance and tube space current, that the heater voltage and current can be adjusted

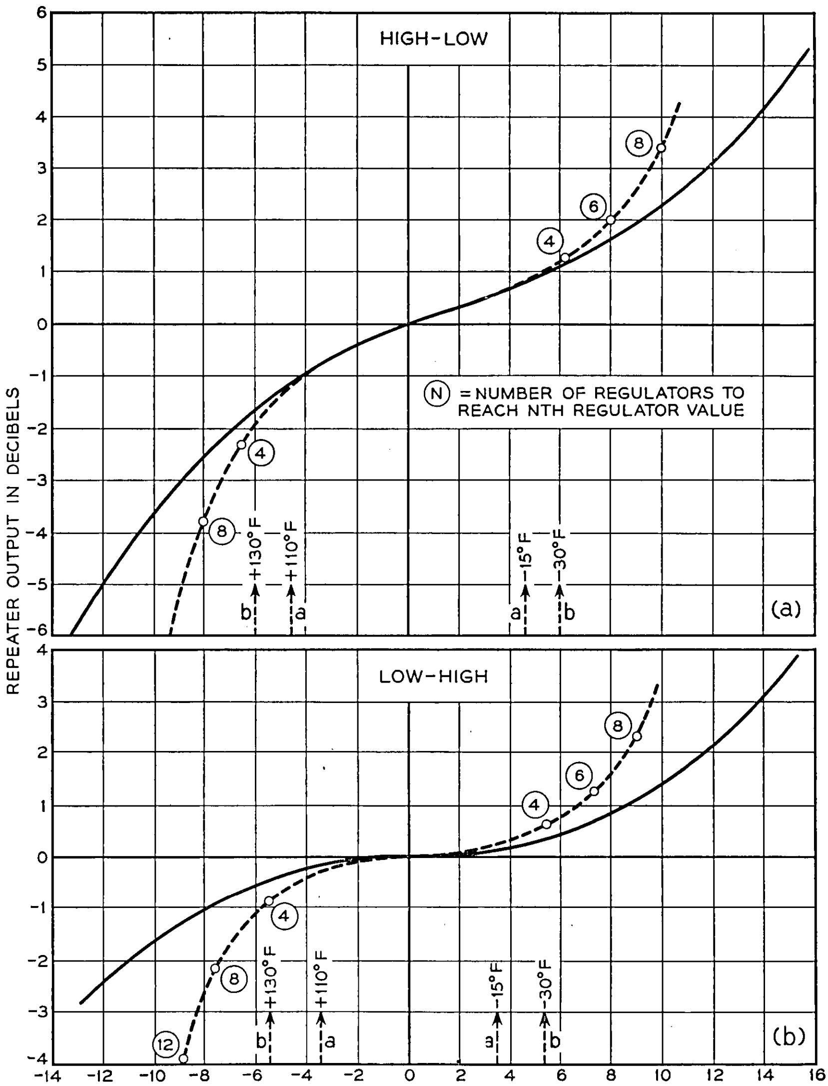


FIG. 13—Regulation characteristics of high-low and low-high repeaters.

to values below those used in telephone offices. As a result an appreciable increase in tube life is obtained compared to that obtained in offices under ordinarily regulated battery conditions.

System Regulation

The regulating characteristics of the LH and HL repeaters are shown in Fig. 13, (a) and (b). The channel unit regulating characteristic is shown in Fig. 14. The solid curves in Fig. 13, (a) and (b), show change in repeater

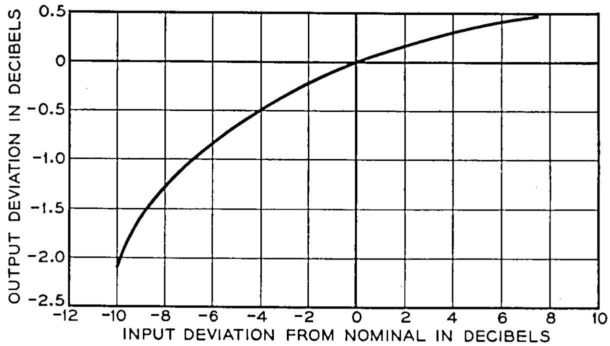


Fig. 14—Regulation characteristic of channel unit regulator.

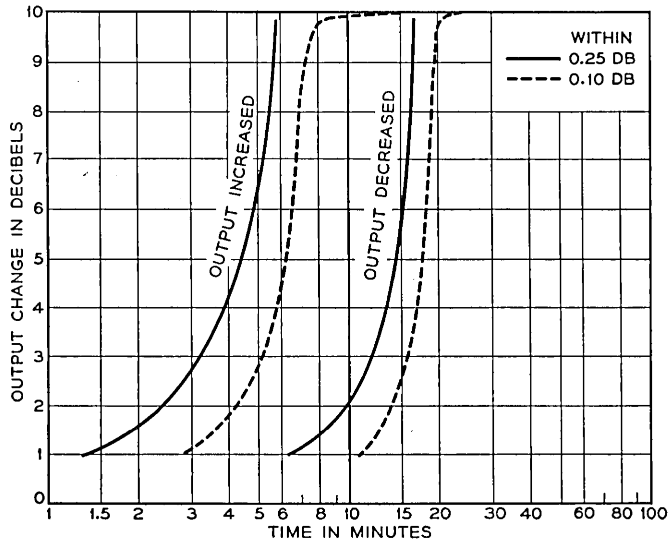


Fig. 15—Stabilization time of repeater or group unit regulator.

output when the total carrier input power is subjected to flat gain line changes as shown by the abscissa. The dotted curves show the regulation of a long string of repeaters, each regulating for the same line change, as well as for the residual output change passed on into successive line sections. The circled numbers indicate the number of the repeater at which the output will depart no further at the indicated input regardless of how many following repeaters there may be. The arrows at "a" and "b" show the line change expected for extreme ranges of ambient temperature for 8 miles of 19-guage toll cable. The group terminal regulators have characteristics about like those of the HL repeater. It can be seen that it would take a most extraordinary set of line conditions to require the channel regulator to compensate for as much as ±5 db change at its input. Despite the doubling of

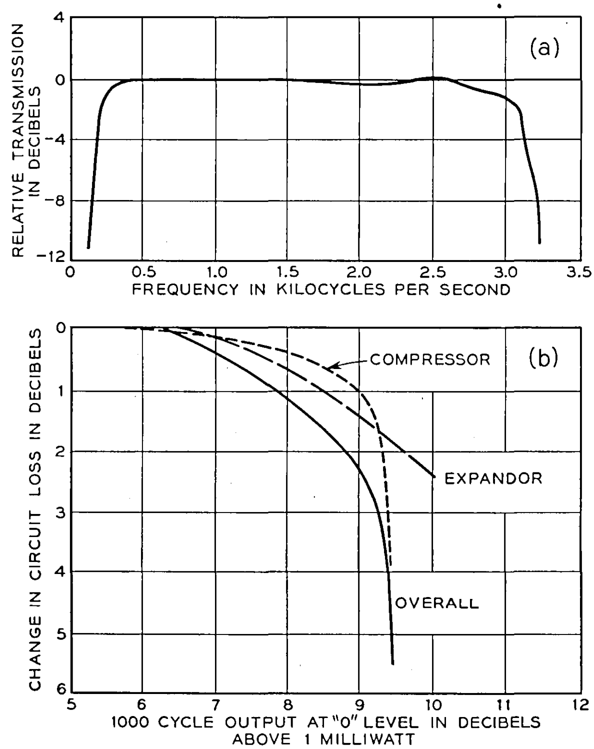


FIG. 16(a)—Typical channel net loss frequency characteristic.
(b)—Typical channel limiting characteristics.

circuit variations by the expandor circuit, Fig. 14 indicates that the overall channel net loss can be expected to stay within about ±2 db.

Figure 15 shows the speed with which the thermistor regulator in the repeater restores the output to normal when subjected to a line change. An increase of input from the line of 6.5 db, for example, requires a 5-minute wait for the output to get within .25 db of its final value. The regulator operates more slowly on decreasing line input. It is essential that the regulator move rather slowly to avoid false regulation on accidental short-duration line hits.

Over-all System Performance

Various means are used to describe the over-all performance of a carrier system such as type N. Subjective tests show, as in other carrier systems, that noticeable deterioration in speech quality occurs when many links are connected in tandem. Satisfactory conversation has been carried on between Milwaukee and Madison, Wisconsin over nine such links, representing a total circuit length of about 750 miles. In this circuit connection, speech passed through 9 compandors, 108 group repeaters and 117 stages of modulation. Practically all of the speech impairment occurred in the 9 compandors.

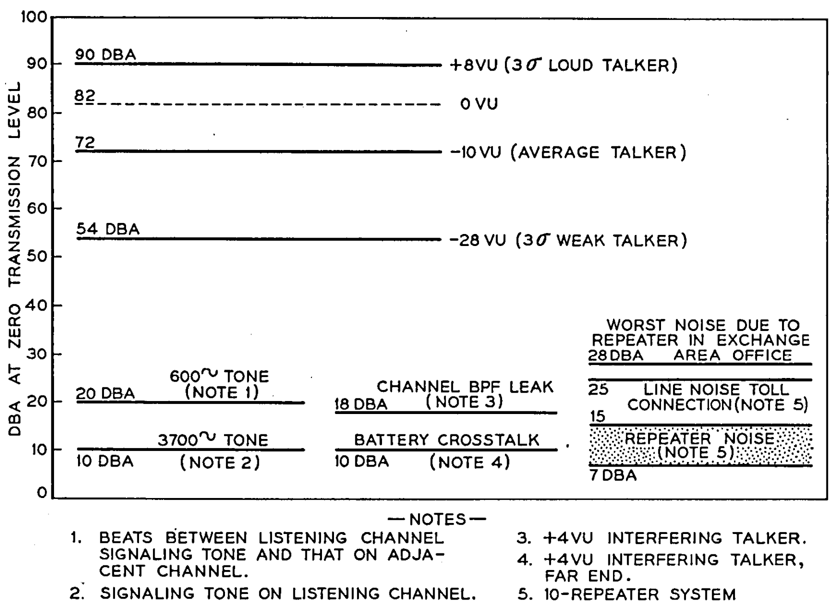


Fig. 17—Relative levels of speech and interference on N-1 carrier.

When only six links were used (which is about the maximum likely to be encountered in service) little impairment was observed. Generally even a critical observer cannot distinguish between a single N channel link and a direct circuit connection between the transmitter and the receiver of the same noise and bandwidth. In Fig. 16, (a) and (b), the frequency characteristic and limiting characteristic of the channel are shown. The useful band of speech frequencies passed is considered to be between 10 db points in four links or about 200 cycles to 3100 cycles. In the N system, because the compandor control circuits are particularly wide-band, the frequency responses are substantially alike when measured with single frequencies or when actuated by speech.

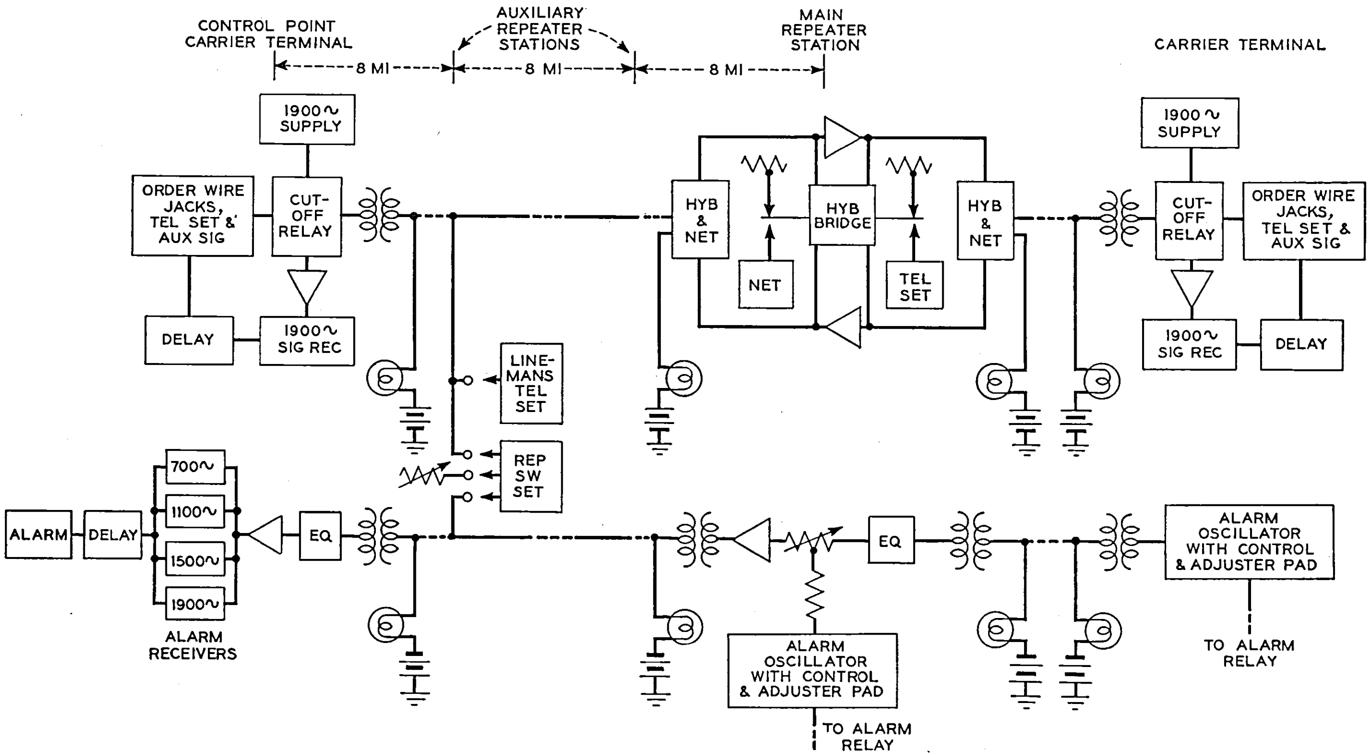


Fig. 18—N-1 order wire and alarm block schematic.

Noise and crosstalk in the over-all system come from many causes. Figure 17 shows relative levels of principal contributions.

Order Wire and Alarm Circuit

Two spare pairs in the cable along an N carrier route are provided for testing and maintenance purposes. One pair, either 16- or 19-gauge with B88 or H172 loading, is used for order wire. Signaling uses either 1900 cycles or 1000–20 ringing. A cableman's whistle is used at pole repeaters to signal attended points. A second pair of conductors is used to bring alarms to attended points from unattended repeater power points. Tones at 700, 1100, 1500 or 1900 cycles are provided for alarming four separate points. Tone is normally on the line and is removed by a relay during a trouble. A 5-second delay in the alarm circuit prevents false operation on line hits. D-C. power is simplexed over the alarm and order wire pairs to pole repeaters as a power source for switching in a spare repeater. Figure 18 shows the order wire and alarm circuit arrangement.

References

1. "Improved Three-Channel Carrier Telephone System," J. T. O'Leary, E. C. Blessing and J. W. Beyer, *Bell Sys. Tech. Jour.*, Jan. 1939.
2. "New Single Channel Carrier Telephone System," H. J. Fisher, M. L. Almquist and R. H. Mills, *A.I.E.E. Trans.* 57, 25, (1938); also *Bell Sys. Tech. Jour.*, Jan. 1938.
3. "Twelve Channel Carrier Telephone System for Open Wire Lines," B. W. Kendall and H. A. Affel, *A.I.E.E. Trans.* 58, 351 (1939); also *Bell Sys. Tech. Jour.*, Jan. 1939.
4. "Some Applications of the Type J. Carrier System," L. C. Starbird and J. D. Mathis, *A.I.E.E. Trans.* 58, 666 (1939); also *Bell Sys. Tech. Jour.*, Apr. 1939.
5. "Carrier Telephone System for Toll Cables," C. W. Green and E. I. Green, *Electrical Engineering*, 57, 227 (1938); also *Bell Sys. Tech. Jour.*, Jan. 1938.
6. "Cable Carrier Telephone Terminals," R. W. Chesnut, L. M. Ilgenfritz and A. Kenner, *Electrical Engineering*, 57, 237 (1938); also *Bell Sys. Tech. Jour.*, Jan. 1938.
7. "Crosstalk and Noise Features of Cable Carrier Telephone Systems," M. A. Weaver, R. S. Tucker and P. S. Darnell, *Electrical Engineering*, 57, 251 (1938); also *Bell Sys. Tech. Jour.*, Jan. 1938.
8. "Experience in Applying Carrier Telephone Systems to Toll Cables," W. B. Bedell, G. B. Ransom and W. A. Stevens, *Bell. Sys. Tech. Jour.*, Oct. 1939.
9. "An Improved Cable Carrier System," H. S. Black and others, *A.I.E.E. Trans.*, 66, 741 (1947).
10. "A Million-cycle Telephone System," M. E. Strieby, *Bell Sys. Tech. Jour.*, Jan. 1937.
11. "Frequency-Division Techniques for a Coaxial Cable Network," R. E. Crane and others, *A.I.E.E. Trans.*, 66, 1451 (1947).
12. "Progress in Coaxial Telephone and Television Systems," L. G. Abraham, *A.I.E.E. Trans.*, 67, 1520 (1984).
13. "Stability of Tandem Regulators with L-1 Carrier Systems," J. P. Kinzer, *A.I.E.E. Trans.* 68 (pt. 2), 1179 (1949).
14. "Attenuation and Delay Equalization for Coaxial Lines," W. R. Lundry, *A.I.E.E. Trans.* 68 (pt. 2), 1174 (1949).
15. "Equalization of Coaxial Lines," K. E. Gould, *A.I.E.E. Trans.* 68 (pt. 2) 1187 (1949).
16. "Carrier Telephone System for Rural Service," J. M. Barstow, *A.I.E.E. Trans.* 66, 501 (1947).
17. "Application of Rural Carrier Telephone System," E. H. Bartelink and others, *A.I.E.E. Trans.* 66, 511 (1947).

Television by Pulse Code Modulation*

By W. M. GOODALL

Transmission by pulse code modulation presents inviting possibilities in the field of television in that information may be relayed by many repeater stations without deterioration. In a PCM system, the information signal is periodically sampled and its instantaneous amplitude described by a group of pulses according to a pre-set code. These pulse groups occur at the sampling rate and constitute the transmitted signal. In this process an operation known as amplitude quantization is required.

This paper will include a discussion of time sampling, amplitude quantization, binary coding and decoding of a television signal. The operation of the equipment used to perform these functions is described.

The results obtained with an experimental system for different numbers of digits (i.e., maximum number of pulses per group) from one to five are illustrated by photographs. The television signal used in these tests was obtained from a special low noise film scanner. As was expected, the number of digits required depends upon the amount of noise in the test signal.

THE papers that have so far appeared on pulse code modulation have dealt primarily with the transmission of speech. The present work deals specifically with the problems involved in the transmission of television, but in its general aspects it is pertinent to the transmission of any broadband signal by PCM. The chief difference between a system for telephony and one for television resides in the required speeds of operation. The use of the wide band required for this system would be justified by the well known advantages of a pulse-code system which have been pointed out by Oliver, Pierce and Shannon[1]. Regenerative repetition of the on and off binary pulses at repeater points permits the relaying of the signal to great distances without introducing any significant degradations due to noise or distortion arising in the medium. In addition, the coding process permits the trading of bandwidth for noise advantage on a very favorable basis.

General Considerations

As is well known, PCM is a form of time-division modulation. The information to be transmitted is sampled at regular intervals. This process results in a definite and limited number of amplitudes per unit of time which replace the original wave in subsequent operations. When the sampling frequency is at least twice the highest frequency present in the original wave, the resulting distortion falls outside the desired band and can be removed by a low-pass filter in the output of the system. For a system of fixed

* Presented orally before the I.R.E. National Convention, New York City, March 1949.

[1] See "Philosophy of PCM"—Oliver, Pierce and Shannon—*Proc. I.R.E.*, Nov. 1948.

sampling frequency it is also desirable to band limit the input signal to avoid undesirable distortion products due to extraneous frequency components which may be present in the original wave. For a nominal 5 mc television channel the sampling rate used in these experiments was 10 mc per second and the input and output filters passed components of 4.3 mc and attenuated components of 5.0 mc. The sampling process produces a discrete number of samples to be transmitted. For the present case this number is 10 million per second.

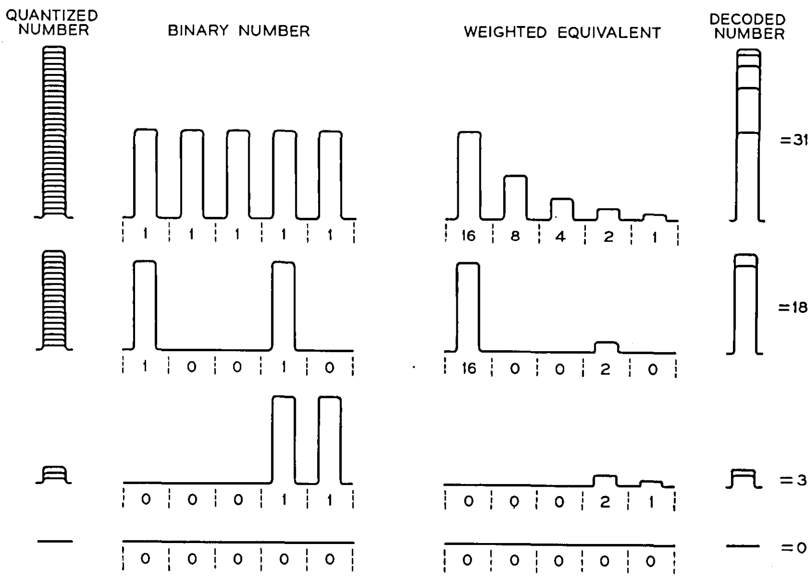


FIG. 1—Five-digit code groups.

Each of these samples may have any value in a continuous range between 0 and a maximum value set by the amplitude range which the system is designed to transmit.

In binary PCM each of these amplitudes is transmitted by a code group of binary digits. As an example, consider a five-digit code which is illustrated in the second column of Fig. 1. Here we have five digits or on-and-off pulses. The maximum number of values that can be represented by these five two-position pulses is 2^5 or 32 values. Examples shown are for amplitudes of 31, 18, 3 and 0. It is easy to see that any other integer value greater than 0 and less than 31 can also be represented by one of the combinations of pulses and spaces. It is also apparent that when all the combinations have been used up no other values can be obtained.

If it were necessary to transmit all of the continuous values present in the sampled wave, it would be necessary to use a large, or even worse, an infinite number of digits. Of course, this is not done. Instead the sampled wave which momentarily may have any value is represented by one of the 32 values that are permitted by the five-digit code. This process is known as amplitude quantization. The quantized amplitudes are shown in the first column of Fig. 1. In the examples shown any number between 17.5 and 18.5 would be represented as 18 and likewise for the other values shown. There is some uncertainty as to the correct value exactly one-half way between permitted values. Here an arbitrary choice is much to be preferred to faulty operation which may give a large error signal. More will be said about this point later.

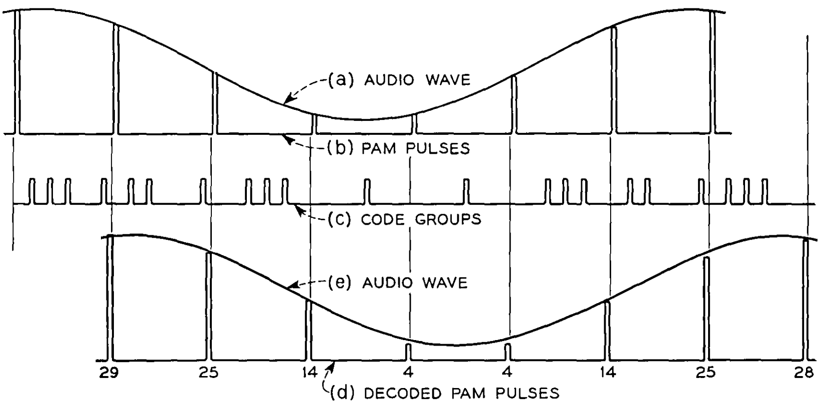


FIG. 2—PCM wave forms.

Each of these code groups, here illustrated as a 5-digit group, represents a single sample of the ten million per second that are needed to represent the television signal. These digits may be transmitted over a single circuit. Figure 2 is an example of this method of transmission for an audio wave *a*. The samples are represented by the PAM wave *b*. The code groups are shown in the wave *c* while the decoded pulses are shown on the wave *d*. The original audio wave, delayed by one sampling interval, is shown as wave *e*. It will be noted that the quantized PAM pulse waves *d* do not fit exactly on the curve. This is the result of the quantization process previously mentioned.

For a five-digit 4 kc telephone channel forty thousand digit pulses per second are used in the transmission medium. For television, the same wave forms apply, and a five-digit 5 mc signal uses fifty million pulses per second in the transmission medium. Figure 2 illustrates a PCM system where the digit pulses are sent on a time-division basis. It is not necessary to do this, however, since the digit pulses may be sent over separate wire or frequency-division carrier circuits. In the experimental setup used in these studies each of the five digits is transmitted over a separate wire circuit. The total bandwidth required in the transmission medium is essentially the same for both methods of transmission. A single one-way television circuit for five digits would require from 50 to 100 megacycles bandwidth in a microwave system. The actual required band would depend upon the state of the art and the complication permitted in the repeater equipment.

From many points of view the transmission medium is the most important part of the system. In non-regenerative systems, for example in the carrier system used in present day coaxial cable transmission, most of the distortion and noise that appears in the final output is the additive resultant of a large number of small contributions arising in the individual repeater links that make up the complete transmission medium. It is easy to see that, for this method of transmission, each repeater link must be much better than the overall system. For a signal that is sampled and quantized in amplitude, however, it is possible to generate a new signal at each repeater which is essentially perfect. In the absence of noise the quantized signal would have one of the permitted amplitudes at the sampling time. A small amount of noise will change this situation so that the amplitude will not be exactly the correct value at the sampling time. As long as the noise or other disturbance is not too great, it is possible to requantize the signal and to transmit the correct amplitude at the sampling times. This process which is known as regeneration can be used for any type of signal that has been sampled and quantized in amplitude. For a system using binary pulses where only two levels are present, the regenerative process is technically possible. Regenerative repeaters would transmit new pulses, which would be accurately timed and properly shaped. As long as the noise is kept below a threshold value, the noise would not accumulate from link to link and the final decoded signal would be of the same quality as one obtained from a monitor located at the transmitter.

This means that the quality of the final output of the system depends upon the size of the time and amplitude quanta used in the PCM system. In other words, the final quality depends upon the sampling rate and the number of digits used and not upon the length of the system.

The last two columns of Fig. 1 show how the digit pulses can be decoded to produce the output signal. The decoder produces the weighted equivalents of the digit pulses which are then added for each code group. Each of these summation pulses represents one of the input samples in a quantized form. These summation pulses are then passed through an appropriate low-pass filter to the output of the system.

Description of Experiments

The experiments to be described were confined to tests with a transmitting terminal connected to the receiving terminal by short coaxial transmission lines. The transmitting terminal performed the functions of sampling, quantizing and coding, while the receiving terminal decoded the PCM signal. No regenerative repeaters were included since they are not necessary in tests designed to evaluate the fundamental limitations of sampling and quantizing of a television signal.

Figure 3 gives a block diagram of the system used in these experiments. The input filter band limits the signal so that the highest frequency is less than one-half of the 10-megacycle sampling rate. The input sampler samples the wave and holds the amplitude value obtained until the next sampling interval. It uses the same type of circuit as that described in the paper by Meacham and Peterson in *The Bell System Technical Journal* for January 1948. In general, much of the circuitry described by them has been used in this equipment, but the units of course function at greatly increased speeds.

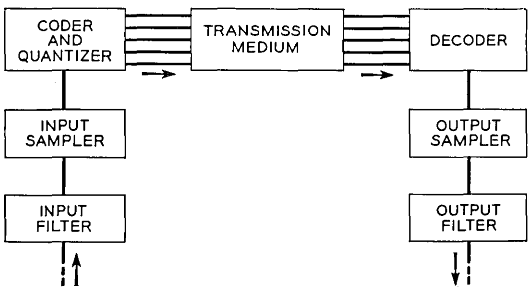


Fig. 3—Block schematic of PCM system.

The coder and quantizer use a coding tube which produces the code simultaneously in five-digit output circuits. Quantization is accomplished by using a special code, together with suitable slicing units in the output of the digit amplifiers. Further discussion of the coder and quantizer will be given in connection with Fig. 4.

In these experiments the transmission medium consisted of an appropriate number of wire circuits, no regenerative repeaters being used. At the receiver, a decoder regenerates the pulses and adds the weighted digits to obtain the quantized PAM signal, as already shown in Figs. 1 and 2. The output sampler is similar to the one used at the input. It will be recalled that step samples are produced, i.e., the signal is sampled at the beginning of each interval and this value is held until the next sampling time. The output filter band limits the signal and removes extraneous components above 5 mc, particularly the 10 mc sampling frequency. As is well known, these filters should have good phase response if they are to be used for a television signal.

The physical equipment used in this experiment is housed in three seven-foot cabinet relay racks. One bay contains the sampler, the push-pull amplifier for driving the deflection plates of the coding tube, the coding tube, the digit amplifier and slicers, and finally the translator. Another bay con-

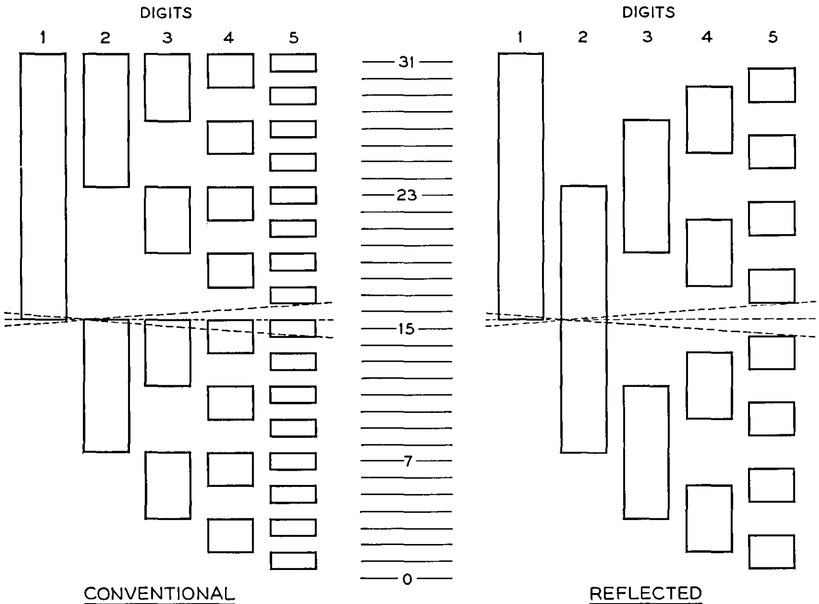


FIG. 4—PCM code plates.

tains the decoder, output sampler, attenuators, patching panel, output filter and other test gear. The last bay contains the regulated power supplies.

CODERS AND QUANTIZERS

We return now to further consideration of the coder and quantizer. The coding tube used in these experiments was developed by Mr. R. W. Sears. It is similar in many respects to one previously described by him in *The Bell System Technical Journal* for January 1948. The older tube produced the code on a time-division basis, while the new tube produces the code simultaneously on a plurality of output digit collectors.

The time-division coding tube which was used by Meacham and Peterson in the 96-channel multiplex system first quantized and then coded the signal. The simultaneous coding tube uses a different code which does not require a quantized input. The encoded signal is subsequently translated into the conventional binary code.

The coder and quantizer are probably the most important parts of the terminals of a PCM system. The following discussion of the two types of coding tubes will illustrate how they function and show how the new tube can function at the greater speeds necessary for television.

Consider the code plate on the left side of the next figure (4). This plate gives the conventional binary code as discussed in connection with the first figure. In a time-division coding tube a point beam is used. It is deflected vertically by the output of the sampler. After the beam has settled down to its proper position, which corresponds to the quantized signal amplitude it is swept across the code plate. An output collector is used which covers the back of the code plate. If the beam goes through a hole in the code plate a pulse is produced; if it is stopped by the code plate no pulse is produced. By this means a code group is produced on a time-division basis for each sample.

As long as the beam does not fall on the edge of a hole, this arrangement functions satisfactorily. Now consider the case where the beam sweeps across the set of edges corresponding to the amplitude 15.5. It is seen that, by a slight misalignment of the code plate and the horizontal deflecting plates, the beam could produce either the code group corresponding to 31 or to 0 depending upon the way the deflection axis is tilted with respect to the code plate. This would result in an error equal to one-half of the total amplitude range of the system. Corresponding errors of smaller magnitude are possible for other levels. In all cases this type of error results for signals which have amplitudes one-half way between values permitted by the code.

Errors of this sort can be avoided by quantization of the signal before the coding. This is accomplished in the earlier tube by using the output from a mesh of grid wires in a feedback arrangement. The wires of the grid overlap the edges of the holes in the code plate. When the beam hits one of these grid wires, a current is fed back into the input which causes the signal amplitude to change in such a way as to move the beam between the grid wires. After a short interval the beam settles down in a quantized condition. Then the beam is swept across the code plate. If the beam tends to become misaligned during the deflection process, the feedback from the grid wires corrects this condition and an accurate code is produced. Because of the time required for the feedback process, this method of quantization limits the number of samples that can be coded in a given time.

Another factor which limits the speed of this type of coder is the time required to sweep the beam across the code plate.

It is apparent that the time required to register the code could be reduced if all digits were produced simultaneously. In this type of coder a line beam is used which covers the full width of the code plate and the code is registered simultaneously on a plurality of digit collectors, one collector being used for each digit of the code. This is one of the features that was included in the coding tube used in these experiments.

In the new tube the time required for quantization by feedback has been avoided by using a different type of code which avoids the large errors which are present in the conventional binary code for amplitude values one-half way between the integer values permitted by the code. This new code, here called the reflected binary code, is shown on the right side of Fig. 4.

For present purposes it should be noted that at points one-half way between integer values the beam intersects only one edge in any code group. Further, if the deflection axis is tilted so that an incorrect code group is indicated, the resulting error is only one quantum level, instead of a much larger value possible with the conventional binary code.

In practice, even if the beam is properly lined up, there will be times when the output for the digit in which the beam intersects an edge will be between zero and full output. Since this digit must be unambiguously represented either by a full pulse or no pulse, it is necessary to make a choice and quantize the particular output under discussion. The output of the digit collectors is amplified and the final quantization of any uncertain digit is done by a slicer which is in the output of each digit amplifier. Use of this code thus localizes final quantization to within a single digit, and an arbitrary choice results at most in an error of one quantum level.

The use of the reflected binary code for PCM applications was suggested to the writer by Mr. F. Gray. As mentioned before, this code is translated to the conventional binary code. The translator used in these tests was designed by Mr. R. L. Carbrey who developed it specifically for this experiment.

Results of Experiments

We now pass on to some of the results obtained with this system. Figure 5 should help in understanding the results shown in the remaining figures. It shows a triangular wave which has been analyzed into the three "on" and "off" rectangular waves shown in the bottom part of the figure.

In this paper we shall follow the convention that the digit of largest amplitude is the first digit, the next largest digit is the second digit and so on. By this convention the first digit is $\frac{1}{2}$ of the total amplitude range, the second digit $\frac{1}{2}$ of the first, or $\frac{1}{4}$ of the total range. Thus, the amplitude of any given digit would be $\frac{1}{2^n}$ of the total amplitude range.

It is convenient to think of the first digit as a first-order approximation to the original in terms of the rectangular waves. The second digit gives a second-order correction to add to the first digit, and the third digit gives a third-order correction to add to the first and second digits.

The rectangular waves, of course, are the envelopes of the pulses that are transmitted over the various digit channels. Because the respective values represented by the various digits are once and for all known, it is not necessary that the amplitudes with which the pulses are transmitted be equal to the values which they represent, but they may to advantage be sent with the same amplitude in all of the digit channels. At the decoder the relative

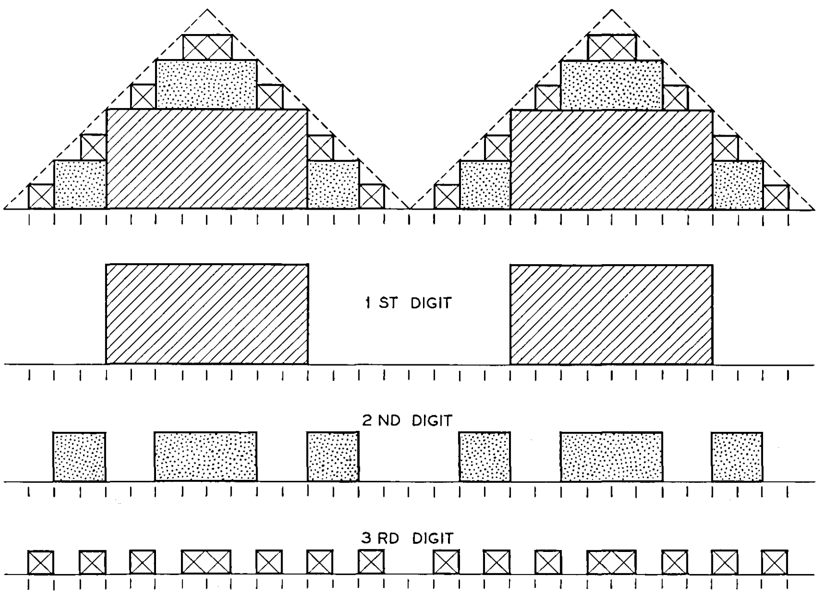


FIG. 5—Rectangular wave approximations.

amplitudes of the digit channels are restored according to the coding convention and the results added to obtain the rectangular wave approximation to the original wave. It is seen that the first three digits give a fair approximation to the original wave. More digits, of course, would improve this approximation.

In general terms, from this point of view, the coder is an analyzer which determines the best approximation to the information wave in terms of a series of rectangular waves of decreasing amplitudes. The decoder is a synthesizer which approximates the original wave by adding the rectangular waves obtained from the coder. The coding convention allows the derived rectangular waves to be transmitted with the same amplitude for all the

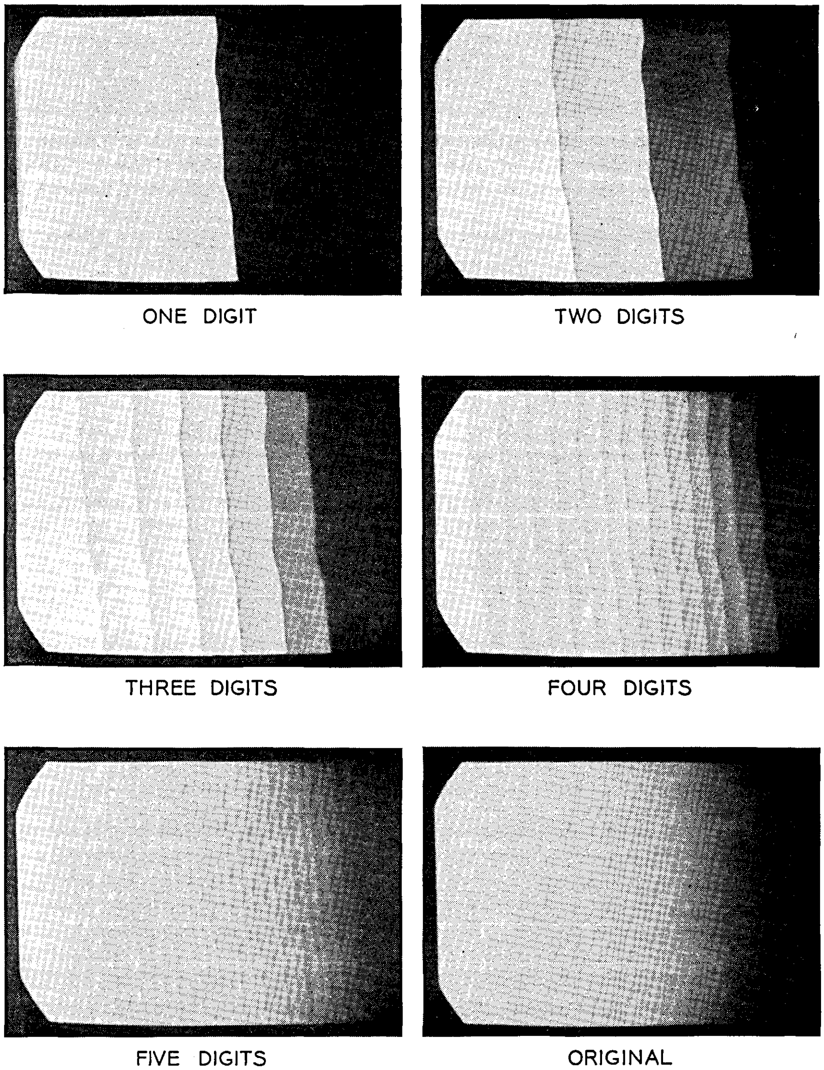


FIG. 6—Test signal, Density wedge: one to five-digit transmission.

digits. The remaining figures show some results obtained by transmitting a television signal through the PCM system described in the first part of the paper.

Mr. B. M. Oliver and others of the television group of The Bell Telephone Laboratories have developed a special low-noise film scanner that provides an excellent test signal. This equipment includes a rooter which, in combination with the expansion of the kinescope, results in an overall linear system. This method of operation, as is well known, results in a wide range of tone values between black and white. The PCM system used, employed steps of equal size; in other words, within the limits of the quantum steps it is a linear system. The combination of the signal from the film scanner, including the rooter, and the linear PCM system, followed by an expanding viewing tube, results in an overall system which employes the limited number of steps in the PCM system to essentially optimum advantage.

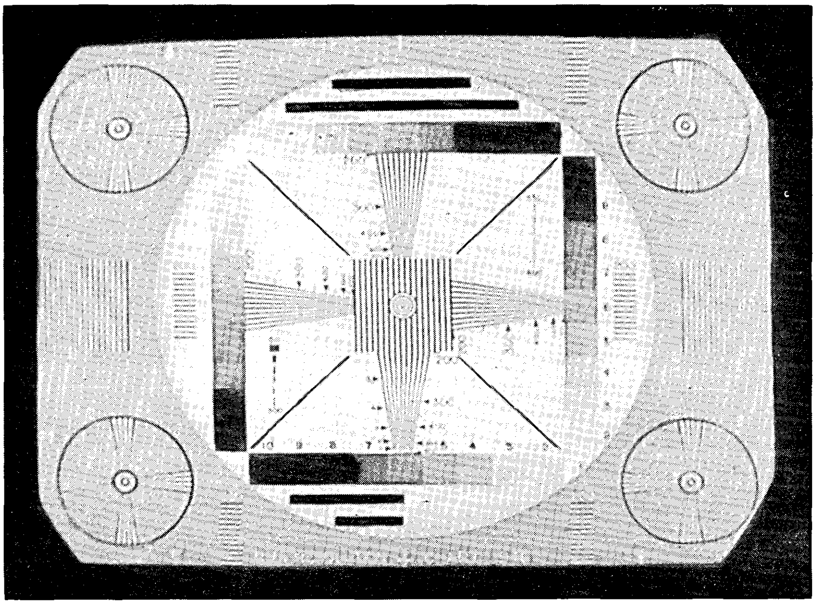


FIG. 7—Test signal, RMA test chart: five-digit transmission.

While in practice, synchronization would probably be derived from the code pulses, for the purposes of this experiment it was not necessary to transmit the synchronizing pulse through the PCM system. Synchronization of the monitor was obtained by a separate path. This was done, since in an operating system it would not be necessary to use more than one or two levels to send the synchronizing pulse as compared with 25% or more of the levels that would be necessary in an unmodified standard television wave-form.

The pictures shown on the figures were taken with a one-second exposure. It will be realized that in a photographic still picture obtained in this manner the exact effect in the viewing tube cannot be conveyed because it is not possible to see motion due to noise.

Figure 6 shows the results obtained using a special test signal for five different PCM systems. The five PCM systems are those which result for

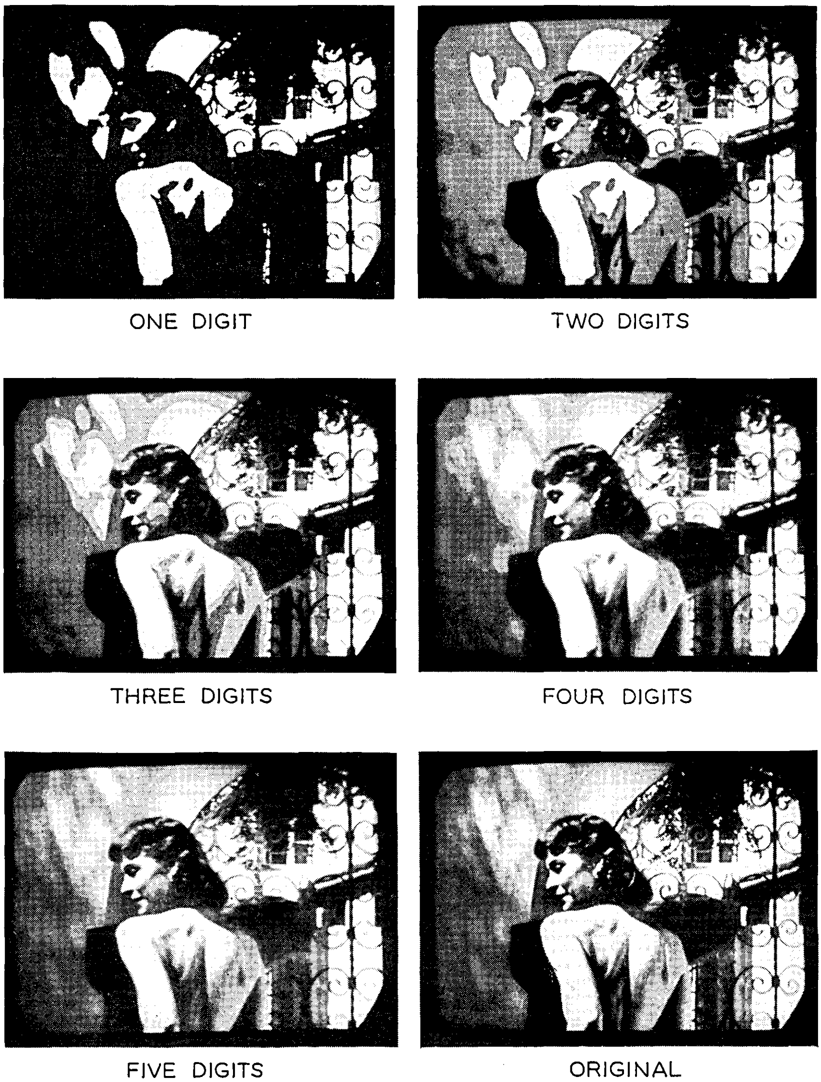
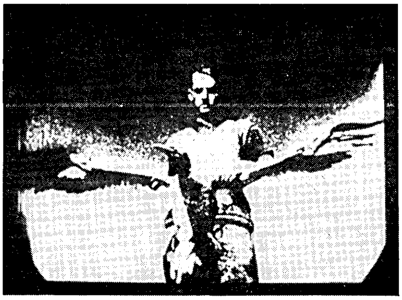
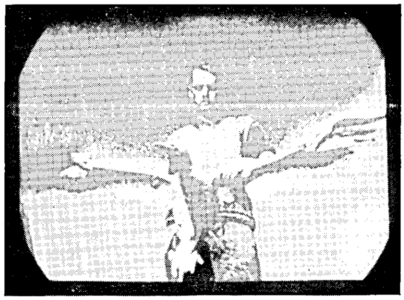


FIG. 8—Test signal, model: one to five-digit transmission.

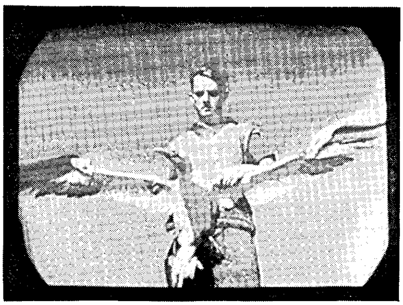
one digit, for two digit, for three, four and five-digit transmission. The test signal for these pictures was an electrical saw tooth wave derived from the horizontal sweep generator. It will be noted that the linear input signal results in an amplitude quantized output signal. The one-digit system sends



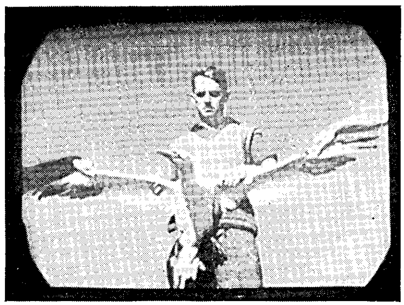
ONE DIGIT



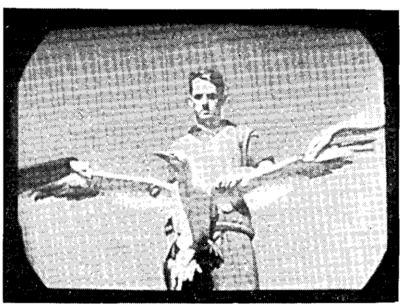
TWO DIGITS



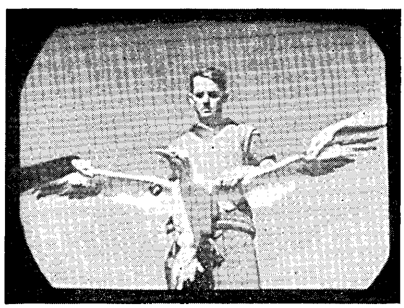
THREE DIGITS



FOUR DIGITS



FIVE DIGITS



ORIGINAL

FIG. 9—Test signal, boy and bird: one to five-digit transmission.

two levels, black and middle grey. The two-digit system sends four levels, the three-digit system sends eight levels and the four and five-digit system

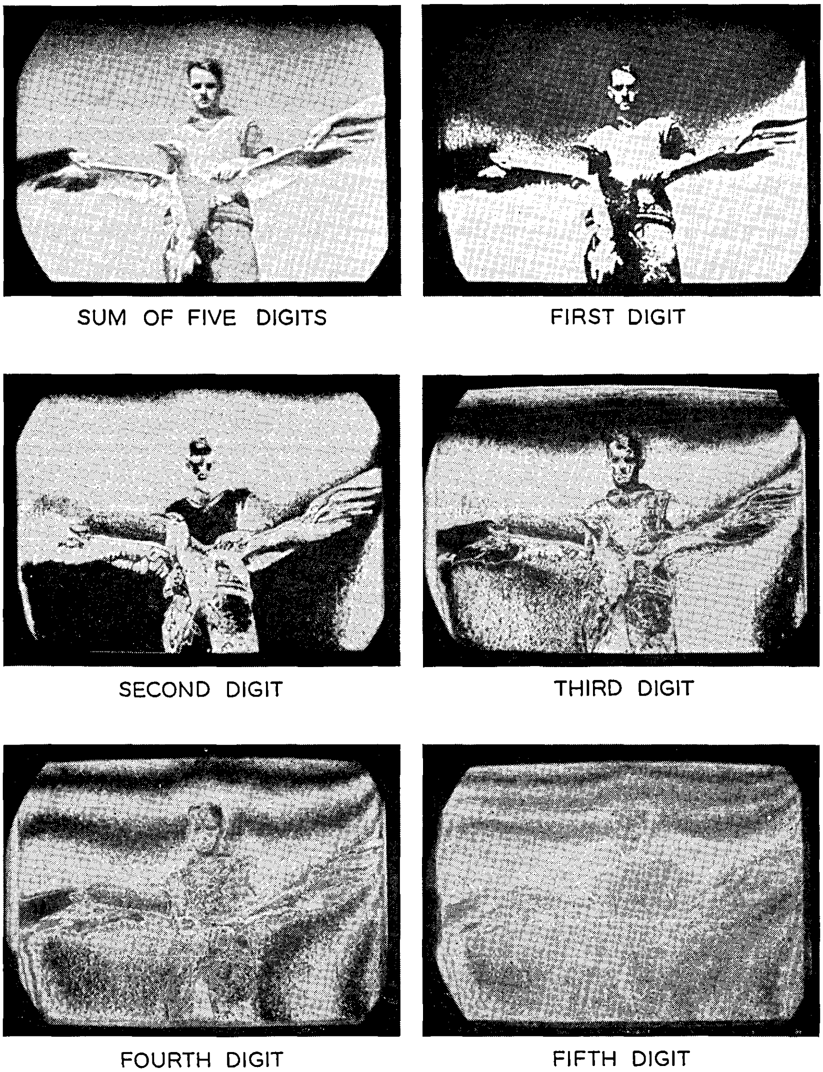


FIG. 10—Test signal, boy and bird: single digits, one through five.

sends sixteen and thirty-two levels. The original, of course, is not quantized and shows a smooth graduation from black to white. It is, however, band limited by the input filter.

The steps in the five-digit system are even more clearly visible on the picture tube. During some tests in which random noise was added to the test signal, it was found that the sharpness of the contour edges was destroyed by random noise when the ratio of the peak-to-peak signal to rms noise was 60 db. For the five-digit picture the smearing of the edges was about one-tenth of the distance between the contours. Other tests which will be described later suggest that the contours for the five-digit thirty-two level system would be masked with an input peak-to-peak signal to rms noise ratio of 40 db.

The writer is not aware of a television system that is capable of generating a signal with a peak-to-peak signal to rms noise ratio of 60 db. However, if such a system were available, these results indicate that an eight- or nine-digit PCM system would be needed to avoid appreciable degradation of the 60 db signal.

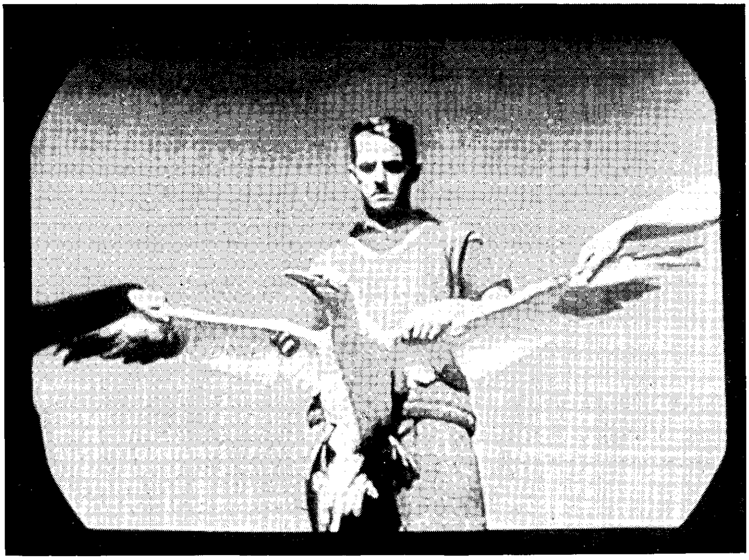
Figure 7 shows the results for an RMA test chart with five-digit PCM transmission. The resolution is limited by the input filter, the film scanner having a resolution corresponding to about 10 megacycles. Using the test pattern for a signal, careful comparison of the band limited transmission with and without the PCM system showed only small defects in the PCM transmission.

When the PCM transmission is seen on the television screen, the contour effects which are strikingly apparent for one, two, and three digits are hardly noticeable for five digits. Figure 8 illustrates this performance as well as is practical with photographic reproduction. About one digit is lost, and the three-digit printed reproduction shows the contours with about the same distinctness as four digits when viewed on the television screen. This statement applies in general to all of the printed reproductions.

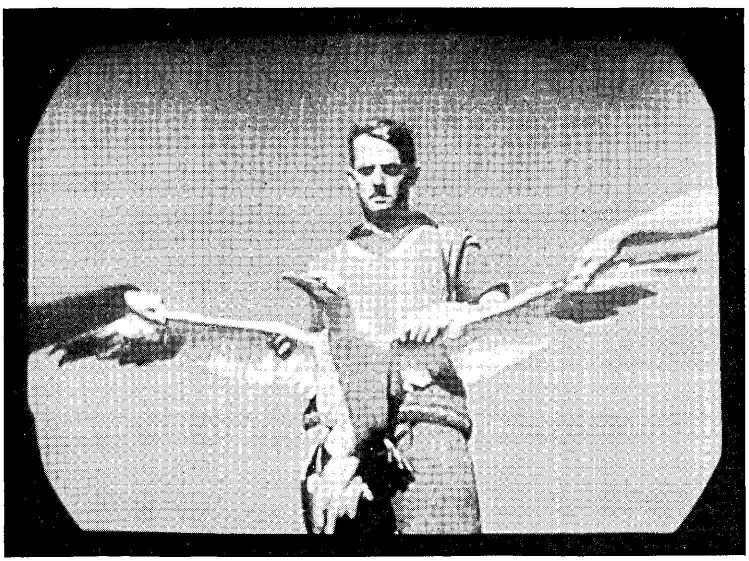
Figure 9 shows the same results for a different subject. The contour effects for a transmission system using a small number of digits are particularly apparent in the sky.

Another method of presenting the results is shown in Fig. 10. In the previous pictures the results have been presented for complete systems using one, two, etc., up to five digits. In Fig. 10, however, the transmission of each of the five digits is separately illustrated. Except for the 5th digit, for which this was not possible, an attempt was made to reproduce the pictures with equal contrast between black and white. The large amount of detail present in the fourth and fifth digits is particularly striking. The sum picture was obtained with proper weighting of the all five digits as discussed earlier in the paper.

The remaining figure (11) illustrates the effect of adding noise to the input to reduce the contour effects. From these photographs it appears that adding noise has been definitely helpful in this respect. However, a penalty is paid for this result. The photographic process reduces the effect of noise by



FOUR DIGITS



FOUR DIGITS PLUS NOISE

FIG. 11—Test signal, boy and bird: with and without added noise.

integration and the picture observed on the monitor for the added noise case was definitely more "noisy" than the other picture. Even so, most observers agreed that in general the adding of noise was desirable for a system using a small number of digits.

ACKNOWLEDGMENTS

Many people contributed to the success of this experiment. Mr. O. E. DeLange and Mr. A. F. Dietrich worked along with the writer in the design, building and testing of the equipment. Mr. R. W. Sears, Mr. R. L. Carbrey and Mr. L. A. Meacham should be specially mentioned in connection with the design of the equipment, and Mr. B. M. Oliver assisted in the television tests. Mr. J. C. Schelleng assisted with suggestions and guidance.

Prediction and Entropy of Printed English

By C. E. SHANNON

(*Manuscript Received Sept. 15, 1950*)

A new method of estimating the entropy and redundancy of a language is described. This method exploits the knowledge of the language statistics possessed by those who speak the language, and depends on experimental results in prediction of the next letter when the preceding text is known. Results of experiments in prediction are given, and some properties of an ideal predictor are developed.

1. Introduction

IN A previous paper[1] the entropy and redundancy of a language have been defined. The entropy is a statistical parameter which measures, in a certain sense, how much information is produced on the average for each letter of a text in the language. If the language is translated into binary digits (0 or 1) in the most efficient way, the entropy H is the average number of binary digits required per letter of the original language. The redundancy, on the other hand, measures the amount of constraint imposed on a text in the language due to its statistical structure, e.g., in English the high frequency of the letter E, the strong tendency of H to follow T or of U to follow Q. It was estimated that when statistical effects extending over not more than eight letters are considered the entropy is roughly 2.3 bits per letter, the redundancy about 50 per cent.

Since then a new method has been found for estimating these quantities, which is more sensitive and takes account of long range statistics, influences extending over phrases, sentences, etc. This method is based on a study of the predictability of English; how well can the next letter of a text be predicted when the preceding N letters are known. The results of some experiments in prediction will be given, and a theoretical analysis of some of the properties of ideal prediction. By combining the experimental and theoretical results it is possible to estimate upper and lower bounds for the entropy and redundancy. From this analysis it appears that, in ordinary literary English, the long range statistical effects (up to 100 letters) reduce the entropy to something of the order of one bit per letter, with a corresponding redundancy of roughly 75%. The redundancy may be still higher when structure extending over paragraphs, chapters, etc. is included. However, as the lengths involved are increased, the parameters in question become more

[1] C. E. Shannon, "A Mathematical Theory of Communication," *Bell System Technical Journal*, v. 27, pp. 379–423, 623–656, July, October, 1948.

erratic and uncertain, and they depend more critically on the type of text involved.

2. Entropy Calculation from the Statistics of English

One method of calculating the entropy H is by a series of approximations $F_0, F_1, F_2, \cdots$, which successively take more and more of the statistics of the language into account and approach H as a limit. F_N may be called the N-gram entropy; it measures the amount of information or entropy due to statistics extending over N adjacent letters of text. F_N is given by[1]

$$\begin{aligned} F_N &= -\sum_{i,j} p(b_i, j) \log_2 p_{b_i}(j) \\ &= -\sum_{i,j} p(b_i, j) \log_2 p(b_i, j) + \sum_i p(b_i) \log p(b_i) \end{aligned} \tag{1}$$

in which: b_i is a block of N-1 letters [(N-1)-gram]

j is an arbitrary letter following b_i

$p(b_i, j)$ is the probability of the N-gram b_i, j

$p_{b_i}(j)$ is the conditional probability of letter j after the block b_i, and is given by $p(b_i, j)/p(b_i)$.

The equation (1) can be interpreted as measuring the average uncertainty (conditional entropy) of the next letter j when the preceding N-1 letters are known. As N is increased, F_N includes longer and longer range statistics and the entropy, H, is given by the limiting value of F_N as $N \to \infty$:

$$H = \lim_{N \to \infty} F_N. \tag{2}$$

The N-gram entropies F_N for small values of N can be calculated from standard tables of letter, digram and trigram frequencies.[2] If spaces and punctuation are ignored we have a twenty-six letter alphabet and F_0 may be taken (by definition) to be $\log_2 26$, or 4.7 bits per letter. F_1 involves letter frequencies and is given by

$$F_1 = -\sum_{i=1}^{26} p(i) \log_2 p(i) = 4.14 \text{ bits per letter.} \tag{3}$$

The digram approximation F_2 gives the result

$$\begin{aligned} F_2 &= -\sum_{i,j} p(i, j) \log_2 p_i(j) \\ &= -\sum_{i,j} p(i, j) \log_2 p(i, j) + \sum_i p(i) \log_2 p(i) \\ &= 7.70 - 4.14 = 3.56 \text{ bits per letter.} \end{aligned} \tag{4}$$

[2] Fletcher Pratt, "Secret and Urgent," Blue Ribbon Books, 1942.

The trigram entropy is given by

$$\begin{aligned} F_3 &= -\sum_{i,j,k} p(i, j, k) \log_2 p_{ij}(k) \\ &= -\sum_{i,j,k} p(i, j, k) \log_2 p(i, j, k) + \sum_{i,j} p(i, j) \log_2 p(i, j) \\ &\doteq 11.0 - 7.7 = 3.3 \end{aligned} \tag{5}$$

In this calculation the trigram table[2] used did not take into account trigrams bridging two words, such as WOW and OWO in TWO WORDS. To compensate partially for this omission, corrected trigram probabilities $p(i, j, k)$ were obtained from the probabilities $p'(i, j, k)$ of the table by the following rough formula:

$$p(i, j, k) = \frac{2.5}{4.5} p'(i, j, k) + \frac{1}{4.5} r(i)p(j, k) + \frac{1}{4.5} p(i, j)s(k)$$

where $r(i)$ is the probability of letter i as the terminal letter of a word and $s(k)$ is the probability of k as an initial letter. Thus the trigrams within words (an average of 2.5 per word) are counted according to the table; the bridging trigrams (one of each type per word) are counted approximately by assuming independence of the terminal letter of one word and the initial digram in the next or vice versa. Because of the approximations involved here, and also because of the fact that the sampling error in identifying probability with sample frequency is more serious, the value of F_3 is less reliable than the previous numbers.

Since tables of N-gram frequencies were not available for $N > 3$, F_4, F_5, etc. could not be calculated in the same way. However, word frequencies have been tabulated[3] and can be used to obtain a further approximation. Figure 1 is a plot on log-log paper of the probabilities of words against frequency rank. The most frequent English word "the" has a probability .071 and this is plotted against 1. The next most frequent word "of" has a probability of .034 and is plotted against 2, etc. Using logarithmic scales both for probability and rank, the curve is approximately a straight line with slope -1; thus, if p_n is the probability of the nth most frequent word, we have, roughly

$$p_n = \frac{.1}{n}. \tag{6}$$

Zipf[4] has pointed out that this type of formula, $p_n = k/n$, gives a rather good approximation to the word probabilities in many different languages. The

[3] G. Dewey, "Relative Frequency of English Speech Sounds," Harvard University Press, 1923.

[4] G. K. Zipf, "Human Behavior and the Principle of Least Effort," Addison-Wesley Press, 1949.

formula (6) clearly cannot hold indefinitely since the total probability Σp_n must be unity, while $\sum_1^\infty .1/n$ is infinite. If we assume (in the absence of any better estimate) that the formula $p_n = .1/n$ holds out to the n at which the

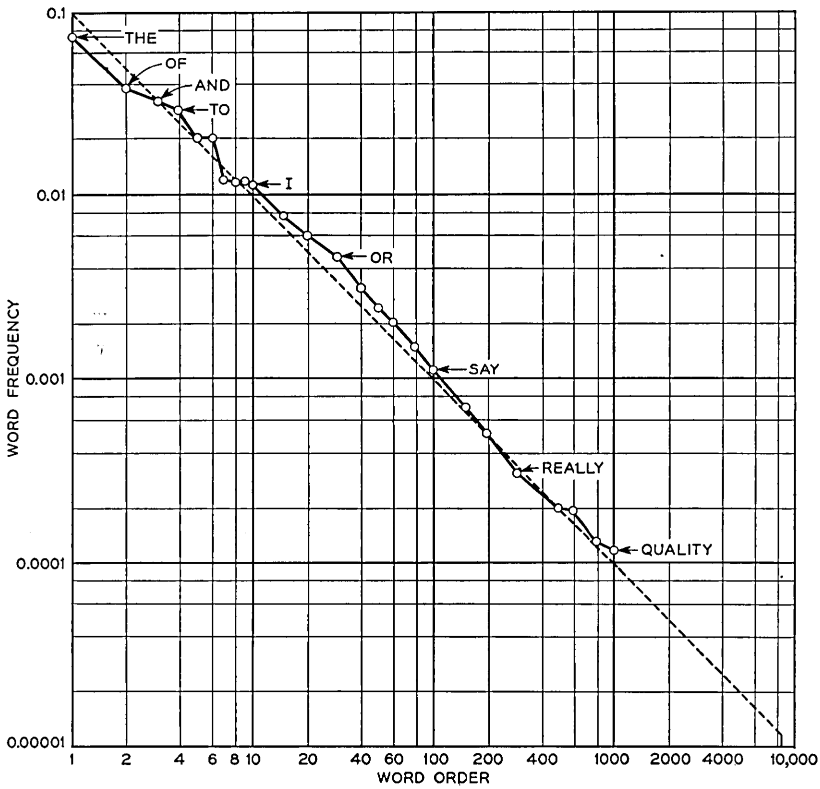


Fig. 1—Relative frequency against rank for English words.

total probability is unity, and that $p_n = 0$ for larger n, we find that the critical n is the word of rank 8,727. The entropy is then:

$$-\sum_1^{8727} p_n \log_2 p_n = 11.82 \text{ bits per word,} \tag{7}$$

or $11.82/4.5 = 2.62$ bits per letter since the average word length in English is 4.5 letters. One might be tempted to identify this value with $F_{4.5}$, but actually the ordinate of the F_N curve at $N = 4.5$ will be above this value. The reason is that F_4 or F_5 involves groups of four or five letters regardless of word division. A word is a cohesive group of letters with strong internal statistical influences, and consequently the N-grams within words are more restricted than those which bridge words. The effect of this is that we have obtained, in 2.62 bits per letter, an estimate which corresponds more nearly to, say, F_5 or F_6.

A similar set of calculations was carried out including the space as an additional letter, giving a 27 letter alphabet. The results of both 26- and 27-letter calculations are summarized below:

	F_0	F_1	F_2	F_3	F_{word}
26 letter..................	4.70	4.14	3.56	3.3	2.62
27 letter..................	4.76	4.03	3.32	3.1	2.14

The estimate of 2.3 for F_8, alluded to above, was found by several methods, one of which is the extrapolation of the 26-letter series above out to that point. Since the space symbol is almost completely redundant when sequences of one or more words are involved, the values of F_N in the 27-letter case will be $\frac{4.5}{5.5}$ or .818 of F_N for the 26-letter alphabet when N is reasonably large.

3. Prediction of English

The new method of estimating entropy exploits the fact that anyone speaking a language possesses, implicitly, an enormous knowledge of the statistics of the language. Familiarity with the words, idioms, clichés and grammar enables him to fill in missing or incorrect letters in proof-reading, or to complete an unfinished phrase in conversation. An experimental demonstration of the extent to which English is predictable can be given as follows: Select a short passage unfamiliar to the person who is to do the predicting. He is then asked to guess the first letter in the passage. If the guess is correct he is so informed, and proceeds to guess the second letter. If not, he is told the correct first letter and proceeds to his next guess. This is continued through the text. As the experiment progresses, the subject writes down the correct text up to the current point for use in predicting future letters. The result of a typical experiment of this type is given below. Spaces were included as an additional letter, making a 27 letter alphabet. The first line is the original text; the second line contains a dash for each letter correctly guessed. In the case of incorrect guesses the correct letter is copied in the second line.

```
(1) THE ROOM WAS NOT VERY LIGHT A SMALL OBLONG              (8)
(2) ----ROO------NOT-V-----I------SM----OBL----
(1) READING LAMP ON THE DESK SHED GLOW ON
(2) REA----------O------D----SHED-GLO--O--
(1) POLISHED WOOD BUT LESS ON THE SHABBY RED CARPET
(2) P-L-S-----O---BU--L-S--O------SH-----RE--C------
```

Of a total of 129 letters, 89 or 69% were guessed correctly. The errors, as would be expected, occur most frequently at the beginning of words and syllables where the line of thought has more possibility of branching out. It might be thought that the second line in (8), which we will call the *reduced text*, contains much less information than the first. Actually, both lines contain the same information in the sense that it is possible, at least in principle, to recover the first line from the second. To accomplish this we need an identical twin of the individual who produced the sequence. The twin (who must be mathematically, not just biologically identical) will respond in the same way when faced with the same problem. Suppose, now, we have only the reduced text of (8). We ask the twin to guess the passage. At each point we will know whether his guess is correct, since he is guessing the same as the first twin and the presence of a dash in the reduced text corresponds to a correct guess. The letters he guesses wrong are also available, so that at each stage he can be supplied with precisely the same information the first twin had available.

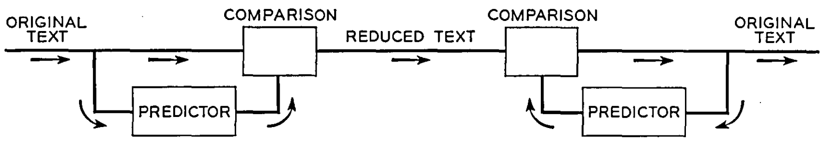


Fig. 2—Communication system using reduced text.

The need for an identical twin in this conceptual experiment can be eliminated as follows. In general, good prediction does not require knowledge of more than N preceding letters of text, with N fairly small. There are only a finite number of possible sequences of N letters. We could ask the subject to guess the next letter for each of these possible N-grams. The complete list of these predictions could then be used both for obtaining the reduced text from the original and for the inverse reconstruction process.

To put this another way, the reduced text can be considered to be an encoded form of the original, the result of passing the original text through a reversible transducer. In fact, a communication system could be constructed in which only the reduced text is transmitted from one point to the other. This could be set up as shown in Fig. 2, with two identical prediction devices.

An extension of the above experiment yields further information concerning the predictability of English. As before, the subject knows the text up to the current point and is asked to guess the next letter. If he is wrong, he is told so and asked to guess again. This is continued until he finds the correct letter. A typical result with this experiment is shown below. The first line is the original text and the numbers in the second line indicate the guess at which the correct letter was obtained.

```
(1) T H E R E   I S   N O   R E V E R S E   O N   A   M O T O R C Y C L E   A
(2) 1 1 1 5 1 1 2 1 1 2 1 1 15 1 17 1 1 1 2 1 3 2 1 2 2 7 1 1 1 1 4 1 1 1 1 1 3 1
(1) F R I E N D   O F   M I N E   F O U N D   T H I S   O U T
(2) 8 6 1 3 1 1 1 1 1 1 1 1 1 1 1 6 2 1 1 1 1 1 1 2 1 1 1 1 1 1
(1) R A T H E R   D R A M A T I C A L L Y   T H E   O T H E R   D A Y
(2) 4 1 1 1 1 1 1 1 1 5 1 1 1 1 1 1 1 1 1 1 1 6 1 1 1 1 1 1 1 1 1 1 1 1 1 1      (9)
```

Out of 102 symbols the subject guessed right on the first guess 79 times, on the second guess 8 times, on the third guess 3 times, the fourth and fifth guesses 2 each and only eight times required more than five guesses. Results of this order are typical of prediction by a good subject with ordinary literary English. Newspaper writing, scientific work and poetry generally lead to somewhat poorer scores.

The reduced text in this case also contains the same information as the original. Again utilizing the identical twin we ask him at each stage to guess as many times as the number given in the reduced text and recover in this way the original. To eliminate the human element here we must ask our subject, for each possible N-gram of text, to guess the most probable next letter, the second most probable next letter, etc. This set of data can then serve both for prediction and recovery.

Just as before, the reduced text can be considered an encoded version of the original. The original language, with an alphabet of 27 symbols, A, B, $\cdots$, Z, space, has been translated into a new language with the alphabet 1, 2, $\cdots$, 27. The translating has been such that the symbol 1 now has an extremely high frequency. The symbols 2, 3, 4 have successively smaller frequencies and the final symbols 20, 21, $\cdots$, 27 occur very rarely. Thus the translating has simplified to a considerable extent the nature of the statistical structure involved. The redundancy which originally appeared in complicated constraints among groups of letters, has, by the translating process, been made explicit to a large extent in the very unequal probabilities of the new symbols. It is this, as will appear later, which enables one to estimate the entropy from these experiments.

In order to determine how predictability depends on the number N of preceding letters known to the subject, a more involved experiment was carried out. One hundred samples of English text were selected at random from a book, each fifteen letters in length. The subject was required to guess the text, letter by letter, for each sample as in the preceding experiment. Thus one hundred samples were obtained in which the subject had available 0, 1, 2, 3, $\cdots$, 14 preceding letters. To aid in prediction the subject made such use as he wished of various statistical tables, letter, digram and trigram tables, a table of the frequencies of initial letters in words, a list of the frequencies of common words and a dictionary. The samples in this experiment were from *"Jefferson the Virginian"* by Dumas Malone. These results, together with a similar test in which 100 letters were known to the subject, are summarized in Table I. The column corresponds to the number of preceding letters known to the subject plus one; the row is the number of the guess. The entry in column N at row S is the number of times the subject guessed the right letter at the Sth guess when (N-1) letters were known. For example,

TABLE I

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	100
1	18.2	29.2	36	47	51	58	48	66	66	67	62	58	66	72	60	80
2	10.7	14.8	20	18	13	19	17	15	13	10	9	14	9	6	18	7
3	8.6	10.0	12	14	8	5	3	5	9	4	7	7	4	9	5	
4	6.7	8.6	7	3	4	1	4	4	4	4	5	6	4	3	5	3
5	6.5	7.1	1	1	3	4	3	6	1	6	5	2	3			4
6	5.8	5.5	4	5	2	3	2			1	4	2	3	4	1	2
7	5.6	4.5	3	3	2	2	8		1	1	1	4	1		4	1
8	5.2	3.6	2	2	1	1	2	1	1	1	1		2	1	3	
9	5.0	3.0	4		5	1	4		2	1	1	2		1		1
10	4.3	2.6	2	1	3		3	1					2			
11	3.1	2.2	2	2	2	1			1	3		1	1	2	1	
12	2.8	1.9	4		2	1	1	1			2	1	1		1	1
13	2.4	1.5	1	1	1	1	1	1	1	1		1	1			
14	2.3	1.2		1			1					1				1
15	2.1	1.0	1	1							1	1	1			
16	2.0	.9					1			1					1	
17	1.6	.7	1		2	1	1				1		2	2		
18	1.6	.5													1	
19	1.6	.4			1	1			1		1					
20	1.3	.3		1		1	1									
21	1.2	.2														
22	.8	.1														
23	.3	.1														
24	.1	.0														
25	.1															
26	.1															
27	.1															

the entry 19 in column 6, row 2, means that with five letters known the cor rect letter was obtained on the second guess nineteen times out of the hun dred. The first two columns of this table were not obtained by the experimental procedure outlined above but were calculated directly from the known letter and digram frequencies. Thus with no known letters the most probable symbol is the space (probability .182); the next guess, if this is wrong, should be E (probability .107), etc. These probabilities are the frequencies with which the right guess would occur at the first, second, etc., trials with best prediction. Similarly, a simple calculation from the digram table gives the entries in column 1 when the subject uses the table to best advantage. Since the frequency tables are determined from long samples of English, these two columns are subject to less sampling error than the others.

It will be seen that the prediction gradually improves, apart from some statistical fluctuation, with increasing knowledge of the past as indicated by the larger numbers of correct first guesses and the smaller numbers of high rank guesses.

One experiment was carried out with "reverse" prediction, in which the subject guessed the letter preceding those already known. Although the task is subjectively much more difficult, the scores were only slightly poorer. Thus, with two 101 letter samples from the same source, the subject obtained the following results:

No. of guess	1	2	3	4	5	6	7	8	>8
Forward	70	10	7	2	2	3	3	0	4
Reverse	66	7	4	4	6	2	1	2	9

Incidentally, the N-gram entropy F_N for a reversed language is equal to that for the forward language as may be seen from the second form in equation (1). Both terms have the same value in the forward and reversed cases.

4. Ideal N-Gram Prediction

The data of Table I can be used to obtain upper and lower bounds to the N-gram entropies F_N. In order to do this, it is necessary first to develop some general results concerning the best possible prediction of a language when the preceding N letters are known. There will be for the language a set of conditional probabilities $p_{i_1, i_2, \cdots, i_{N-1}}(j)$. This is the probability when the (N-1) gram $i_1, i_2, \cdots, i_{N-1}$ occurs that the next letter will be j. The best guess for the next letter, when this (N-1) gram is known to have occurred, will be that letter having the highest conditional probability. The second guess should be that with the second highest probability, etc. A machine or person guessing in the best way would guess letters in the order of decreasing conditional probability. Thus the process of reducing a text with such an ideal predictor consists of a mapping of the letters into the numbers from 1 to 27 in such a way that the most probable next letter [conditional on the known preceding (N-1) gram] is mapped into 1, etc. The frequency of 1's in the reduced text will then be given by

$$q_1^N = \Sigma p(i_1, i_2, \cdots, i_{N-1}, j) \tag{10}$$

where the sum is taken over all (N-1) grams $i_1, i_2, \cdots, i_{N-1}$ the j being the one which maximizes p for that particular (N-1) gram. Similarly, the frequency of 2's, q_2^N, is given by the same formula with j chosen to be that letter having the second highest value of p, etc.

On the basis of N-grams, a different set of probabilities for the symbols in the reduced text, $q_1^{N+1}, q_2^{N+1}, \ldots, q_{27}^{N+1}$, would normally result. Since this prediction is on the basis of a greater knowledge of the past, one would expect the probabilities of low numbers to be greater, and in fact one can prove the following inequalities:

$$\sum_{i=1}^{S} q_i^{N+1} \geq \sum_{i=1}^{S} q_i^N \qquad S = 1, 2, \cdots . \tag{11}$$

This means that the probability of being right in the first S guesses when the preceding N letters are known is greater than or equal to that when only (N-1) are known, for all S. To prove this, imagine the probabilities $p(i_1, i_2, \cdots, i_N, j)$ arranged in a table with j running horizontally and all the N-grams vertically. The table will therefore have 27 columns and 27^N rows. The term on the left of (11) is the sum of the S largest entries in each row, summed over all the rows. The right-hand member of (11) is also a sum of entries from this table in which S entries are taken from each row but not necessarily the S largest. This follows from the fact that the right-hand member would be calculated from a similar table with (N-1) grams rather than N-grams listed vertically. Each row in the N-1 gram table is the sum of 27 rows of the N-gram table, since:

$$p(i_2, i_3, \cdots, i_N, j) = \sum_{i_1=1}^{27} p(i_1, i_2, \cdots, i_N, j). \tag{12}$$

The sum of the S largest entries in a row of the N-1 gram table will equal the sum of the $27S$ selected entries from the corresponding 27 rows of the N-gram table only if the latter fall into S columns. For the equality in (11) to hold for a particular S, this must be true of every row of the N-1 gram table. In this case, the first letter of the N-gram does not affect the set of the S most probable choices for the next letter, although the ordering within the set may be affected. However, if the equality in (11) holds for all S, it follows that the ordering as well will be unaffected by the first letter of the N-gram. The reduced text obtained from an ideal N-1 gram predictor is then identical with that obtained from an ideal N-gram predictor.

Since the partial sums

$$Q_S^N = \sum_{i=1}^{S} q_i^N \qquad S = 1, 2, \cdots \tag{13}$$

are monotonic increasing functions of N, <1 for all N, they must all approach limits as $N \to \infty$. Their first differences must therefore approach limits as $N \to \infty$, i.e., the q_i^N approach limits, q_i^∞. These may be interpreted as the relative frequency of correct first, second, $\cdots$, guesses with knowledge of the entire (infinite) past history of the text.

The ideal N-gram predictor can be considered, as has been pointed out, to be a transducer which operates on the language translating it into a sequence of numbers running from 1 to 27. As such it has the following two properties:

1. The output symbol is a function of the present input (the predicted next letter when we think of it as a predicting device) and the preceding (N-1) letters.
2. It is *instantaneously* reversible. The original input can be recovered by a suitable operation on the reduced text without loss of time. In fact, the inverse operation also operates on only the (N-1) preceding symbols of the reduced text together with the present output.

The above proof that the frequencies of output symbols with an N-1 gram predictor satisfy the inequalities:

$$\sum_1^S q_i^N \geq \sum_1^S q_i^{N-1} \qquad S = 1, 2, \cdots, 27 \tag{14}$$

can be applied to any transducer having the two properties listed above. In fact we can imagine again an array with the various (N-1) grams listed vertically and the present input letter horizontally. Since the present output is a function of only these quantities there will be a definite output symbol which may be entered at the corresponding intersection of row and column. Furthermore, the instantaneous reversibility requires that no two entries in the same row be the same. Otherwise, there would be ambiguity between the two or more possible present input letters when reversing the translation. The total probability of the S most probable symbols in the output, say $\sum_1^S r_i$, will be the sum of the probabilities for S entries in each row, summed over the rows, and consequently is certainly not greater than the sum of the S largest entries in each row. Thus we will have

$$\sum_1^S q_i^N \geq \sum_1^S r_i \qquad S = 1, 2, \cdots, 27 \tag{15}$$

In other words ideal prediction as defined above enjoys a preferred position among all translating operations that may be applied to a language and which satisfy the two properties above. Roughly speaking, ideal prediction collapses the probabilities of various symbols to a small group more than any other translating operation involving the same number of letters which is instantaneously reversible.

Sets of numbers satisfying the inequalities (15) have been studied by Muirhead in connection with the theory of algebraic inequalities.[5] If (15) holds when the q_i^N and r_i are arranged in decreasing order of magnitude, and

[5] Hardy, Littlewood and Polya, "Inequalities," Cambridge University Press, 1934.

also $\sum_1^{27} q_i^N = \sum_1^{27} r_i$, (this is true here since the total probability in each case is 1), then the first set, q_i^N, is said to *majorize* the second set, r_i. It is known that the majorizing property is equivalent to either of the following properties:

1. The r_i can be obtained from the q_i^N by a finite series of "flows." By a flow is understood a transfer of probability from a larger q to a smaller one, as heat flows from hotter to cooler bodies but not in the reverse direction.
2. The r_i can be obtained from the q_i^N by a generalized "averaging" operation. There exists a set of non-negative real numbers, a_{ij}, with $\sum_j a_{ij} = \sum_i a_{ij} = 1$ and such that

$$r_i = \sum_j a_{ij} (q_j^N). \tag{16}$$

5. Entropy Bounds from Prediction Frequencies

If we know the frequencies of symbols in the reduced text with the ideal N-gram predictor, q_i^N, it is possible to set both upper and lower bounds to the N-gram entropy, F_N, of the original language. These bounds are as follows:

$$\sum_{i=1}^{27} i(q_i^N - q_{i+1}^N) \log i \leq F_N \leq - \sum_{i=1}^{27} q_i^N \log q_i^N. \tag{17}$$

The upper bound follows immediately from the fact that the maximum possible entropy in a language with letter frequencies q_i^N is $-\sum q_i^N \log q_i^N$. Thus the entropy per symbol of the reduced text is not greater than this. The N-gram entropy of the reduced text is equal to that for the original language, as may be seen by an inspection of the definition (1) of F_N. The sums involved will contain precisely the same terms although, perhaps, in a different order. This upper bound is clearly valid, whether or not the prediction is ideal.

The lower bound is more difficult to establish. It is necessary to show that with any selection of N-gram probabilities $p(i_1, i_2, \ldots, i_N)$, we will have

$$\sum_{i=1}^{27} i(q_i^N - q_{i+1}^N) \log i \leq \sum_{i_1, \cdots, i_N} p(i, \cdots i_N) \log p_i, \cdots i_{N-1}(i_N) \tag{18}$$

The left-hand member of the inequality can be interpreted as follows: Imagine the q_i^N arranged as a sequence of lines of decreasing height (Fig. 3). The actual q_i^N can be considered as the sum of a set of rectangular distributions as shown. The left member of (18) is the entropy of this set of distributions. Thus, the i^{th} rectangular distribution has a total probability of

$i(q_i^N - q_{i+1}^N)$. The entropy of the distribution is log i. The total entropy is then

$$\sum_{i=1}^{27} i(q_i^N - q_{i+1}^N) \log i.$$

The problem, then, is to show that any system of probabilities $p(i_1, \ldots, i_N)$, with best prediction frequencies q_i has an entropy F_N greater than or equal to that of this rectangular system, derived from the same set of q_i.

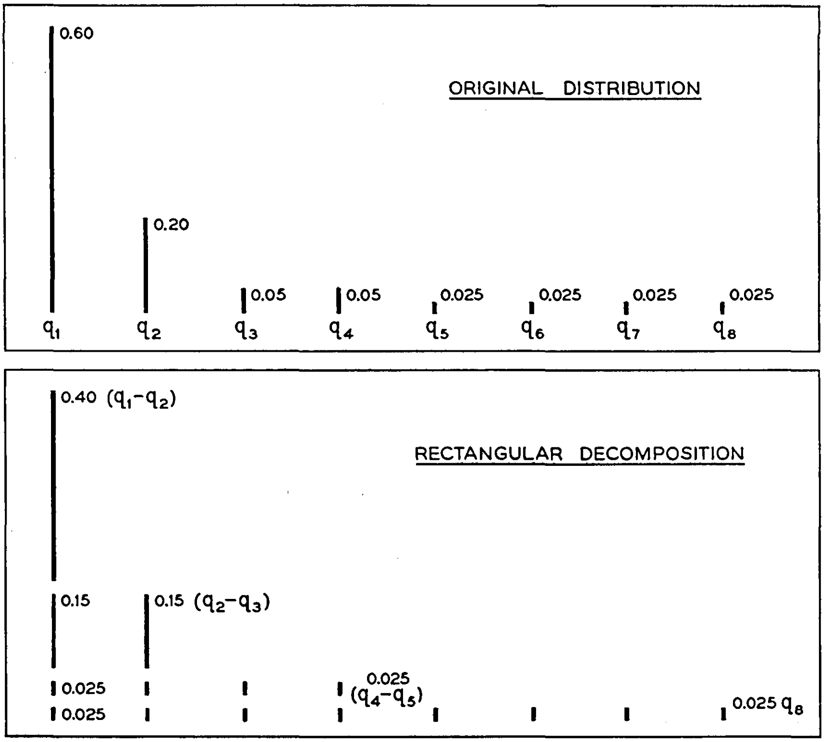


Fig. 3—Rectangular decomposition of a monotonic distribution.

The q_i as we have said are obtained from the $p(i_1, \ldots, i_N)$ by arranging each row of the table in decreasing order of magnitude and adding vertically. Thus the q_i are the sum of a set of monotonic decreasing distributions. Replace each of these distributions by its rectangular decomposition. Each one is replaced then (in general) by 27 rectangular distributions; the q_i are the sum of 27 x 27^N rectangular distributions, of from 1 to 27 elements, and all starting at the left column. The entropy for this set is less than or equal to that of the original set of distributions since a termwise addition of two or more distributions always increases entropy. This is actually an application of the general theorem that $H_y(x) \leq H(x)$ for any chance variables x and y. The equality holds only if the distributions being added are proportional. Now we may add the different components of the same width without changing the entropy (since in this case the distributions *are* proportional). The result is that we have arrived at the rectangular decomposition of the q_i, by a series of processes which decrease or leave constant the entropy, starting with the original N-gram probabilities. Consequently the entropy of the original system F_N is greater than or equal to that of the rectangular decomposition of the q_i. This proves the desired result.

It will be noted that the lower bound is definitely less than F_N unless each row of the table has a rectangular distribution. This requires that for each

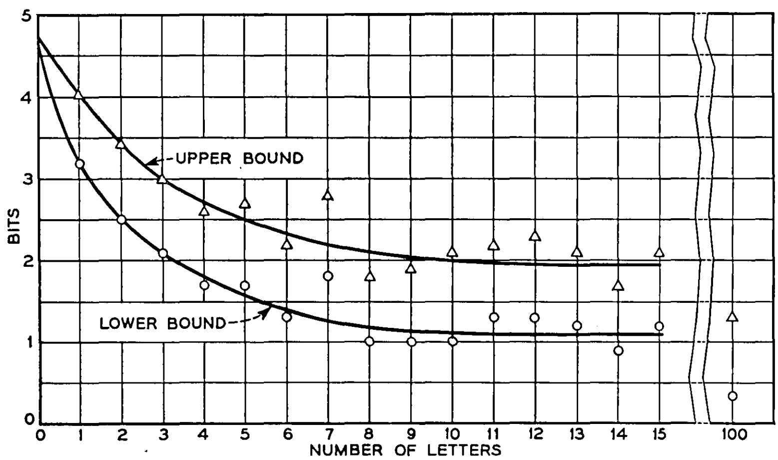


Fig. 4—Upper and lower experimental bounds for the entropy of 27-letter English.

possible (N-1) gram there is a set of possible next letters each with equal probability, while all other next letters have zero probability.

It will now be shown that the upper and lower bounds for F_N given by (17) are monotonic decreasing functions of N. This is true of the upper bound since the q_i^{N+1} majorize the q_i^N and any equalizing flow in a set of probabilities increases the entropy. To prove that the lower bound is also monotonic decreasing we will show that the quantity

$$U = \sum_i i(q_i - q_{i+1}) \log i \tag{20}$$

is increased by an equalizing flow among the q_i. Suppose a flow occurs from q_i to q_{i+1}, the first decreased by Δq and the latter increased by the same amount. Then three terms in the sum change and the change in U is given by

$$\Delta U = [-(i-1) \log (i-1) + 2i \log i - (i+1) \log (i+1)]\Delta q \tag{21}$$

The term in brackets has the form $-f(x-1) + 2f(x) - f(x+1)$ where $f(x) = x \log x$. Now $f(x)$ is a function which is concave upward for positive x, since $f''(x) = 1/x > 0$. The bracketed term is twice the difference between the ordinate of the curve at $x = i$ and the ordinate of the midpoint of the chord joining $i - 1$ and $i + 1$, and consequently is negative. Since Δq also is negative, the change in U brought about by the flow is positive. An even simpler calculation shows that this is also true for a flow from q_1 to q_2 or from q_{26} to q_{27} (where only two terms of the sum are affected). It follows that the lower bound based on the N-gram prediction frequencies q_i^N is greater than or equal to that calculated from the $N + 1$ gram frequencies q_i^{N+1}.

6. Experimental Bounds for English

Working from the data of Table I, the upper and lower bounds were calculated from relations (17). The data were first smoothed somewhat to overcome the worst sampling fluctuations. The low numbers in this table are the least reliable and these were averaged together in groups. Thus, in column 4, the 47, 18 and 14 were not changed but the remaining group totaling 21 was divided uniformly over the rows from 4 to 20. The upper and lower bounds given by (17) were then calculated for each column giving the following results:

Column	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	100
Upper.......	4.03	3.42	3.0	2.6	2.7	2.2	2.8	1.8	1.9	2.1	2.2	2.3	2.1	1.7	2.1	1.3
Lower.......	3.19	2.50	2.1	1.7	1.7	1.3	1.8	1.0	1.0	1.0	1.3	1.3	1.2	.9	1.2	.6

It is evident that there is still considerable sampling error in these figures due to identifying the observed sample frequencies with the prediction probabilities. It must also be remembered that the lower bound was proved only for the ideal predictor, while the frequencies used here are from human prediction. Some rough calculations, however, indicate that the discrepancy between the actual F_N and the lower bound with ideal prediction (due to the failure to have rectangular distributions of conditional probability) more than compensates for the failure of human subjects to predict in the ideal manner. Thus we feel reasonably confident of both bounds apart from sampling errors. The values given above are plotted against N in Fig. 4.

Acknowledgment

The writer is indebted to Mrs. Mary E. Shannon and to Dr. B. M. Oliver for help with the experimental work and for a number of suggestions and criticisms concerning the theoretical aspects of this paper.

A Submarine Telephone Cable with Submerged Repeaters

By J. J. GILBERT

(*Manuscript Received Sept. 11, 1950*)

The paper describes the recently installed Key West-Havana submarine cable telephone system in which repeaters designed for long life are incorporated in the cable structure and are laid as part of the cable.

IN APRIL of last year there was installed between Key West, Florida, and Havana, Cuba, a submarine telephone cable system involving a radical departure from the conventional art of long distance submarine telephony. This departure consisted of the inclusion within the armor of the submarine cable of electron tube repeaters which are designed to pass through the cable laying machinery and sink to the ocean bottom like a length of cable, and which, over an extended period of perhaps twenty years, should not require servicing for the purpose of changing electron tubes or defective circuit elements. The repeater has the appearance of a bulge in the cable about three inches in diameter and tapering off in both directions to the cable diameter of a little over an inch. The total length of the bulge including the taper at each end is about 35 feet. The bulge is flexible enough so that it can conform to the curvature of the brake drum and of the various sheaves in the laying gear on the cable ship. A repeater, with stub cables, is shown in Fig. 1.

Historical

The new cable system, comprising cables Nos. 5 and 6 of the Cuban-American Telephone and Telegraph Company, represents another step in the development of telephonic communication between the United States and Cuba, which has presented many interesting problems. Natural conditions make it difficult, if not impossible, to employ some of the usual methods of communication. One such condition is the absence of high ground in Florida that would permit the use of economic radio systems. Another is the stretch of water between Florida and Cuba, which, in places, is as much as 6,000 feet in depth and which restricts the type of cable that can be used. The practical solution has been to go from the point of contact with the Bell System toll lines at Miami, over the Keys to Key West by land line (with some water crossings), thence to Havana, an air line distance of about 100 n.m., by submarine cable of the deep sea type of construction, having a single coaxial circuit, insulated with water resistant material. There are problems involved in building and maintaining the Miami-Key West connections but these are outside the scope of the present article.

Telephone communication between the United States and Cuba was initiated in 1921, when three submarine cables were laid between Key West and Havana.[1] Each cable provided a telephone circuit, operated on a two-wire basis, and two or more telegraph circuits, d-c. and a-c. The cables were continuously loaded with iron wire, insulated with gutta-percha and had return conductors consisting of copper tapes laid on the insulated core and exposed

Fig. 1—Submarine cable repeater.

electrically to the sea water. These cables were the first ones to employ the copper return conductor, which has also been used in subsequent cables. The copper return was employed after a theoretical study had indicated that the armor and sea water, which for the low frequencies then involved in cable telegraphy furnished a low resistance return conductor, would not be satisfactory at voice and higher frequencies. At these frequencies skin effect causes the return current to concentrate in the armor wires, which are naturally poor

[1]W. H. Martin, G. A. Anderegg, B. W. Kendall, Key West-Havana Submarine Telephone Cable System, *A.I.E.E. Transactions*, Vol. 41, pp. 1–19, February 1922.

conductors for alternating currents, and this makes the resistance of the return path rise to undesirable values. The copper return effectively removes the armor wire and sea water from the transmission circuit at all but very low frequencies. This has the further advantage of reducing the exposure of the circuit to static noise.

Although the iron wire loading was very effective in reducing attenuation in the voice range, the eddy current resistance due to the loading wire made itself felt to a rapidly increasing degree for frequencies above the voice band. Consequently, when additional circuits were required some years later and it was decided to extend the frequency range in order to make use of newly developed carrier frequency equipment, it was necessary to dispense with magnetic loading. In 1930 the Key West-Havana No. 4 cable was laid embodying new materials and novel principles of design.[2] The insulating material in this case was paragutta, which had been recently developed by the Laboratories, and which possessed electrical characteristics and stability much superior to gutta-percha. An intensive study had been made of the design of coaxial cables for carrier frequencies with the aim of obtaining optimum electrical performance by proper proportioning and construction of the conductors and these principles were employed in the new cable. Initially, three carrier telephone circuits were obtained on the No. 4 cable using the equivalent four-wire method, with separate frequency bands for transmission in opposite directions. The cable had been designed with considerable transmission margin and in 1940 the need for additional circuits to Cuba resulted in the installation of new terminal equipment which enabled it to provide seven two-way high quality circuits on an equivalent four-wire basis.

The Key West-Havana No. 4 cable design has proved very popular in other parts of the world. Several such cables have been laid between England and the Continent and between England and Ireland, between Australia and Tasmania, and others were used in connection with the war effort. A cable of this design has also been laid between two of the Japanese islands.

Submerged Repeaters

The demands for circuits to Havana continued to grow, and, after the close of World War II, the time appeared ripe for making use of a new development which had just come to a head in the Laboratories after a period of experimentation.[3] This development was the submerged repeater. The need for periodic strengthening of signals transmitted over considerable distances is about the same in submarine cables as it is in land

[2] H. A. Affel, W. S. Gorton, R. W. Chesnut, New Key West-Havana Carrier Telephone Cable, *B.S.T.J.*, Vol. 11, pp. 197–212, April 1932.

[3] O. E. Buckley, The Future of Transoceanic Telephony, Thirty-Third Kelvin Lecture, *B.S.T.J.*, Vol. 21, pp. 1–19, July 1942.

lines and, as in land lines, the permissible spacing between repeaters usually diminishes as the desired frequency increases. The great difficulty in the case of submarine cable routes is that there are usually no land sites on which repeaters can be located. Artificial islands consisting of floating platforms or buoys have been proposed as a solution, but ocean currents and storms have disastrous effects on such structures. Interruptions due to such causes would make it difficult to meet the requirement on continuity of service which is necessary in the case of important telephone circuits. Therefore, it appeared that the safest place for a submarine cable repeater is on the ocean bottom.

Requirements on Repeater

The decision to place the repeater on the ocean bottom resulted in special requirements on the structure the first of which is that it be capable of resisting the considerable hydrostatic pressure that is encountered in deep water. It also seemed desirable that the operation of getting the repeater overboard from the cable ship should not impede the smooth functioning of the laying process. The best way of meeting this requirement appeared to be to make the repeater structure flexible, within practicable limits, and as small as possible in diameter so that it could pass around the drum and sheaves of the laying gear like any length of cable.

In order to make such a repeater attractive from operating and commercial points of view another requirement was necessary, namely, that the electrical circuit elements of the repeater, including electron tubes, resistances, condensers and coils be designed for long life under operating conditions, so that there would be assurance of freedom from trouble or need of replacement of parts over a long period, perhaps twenty years or more. Servicing of the repeater would be in the nature of a cable repair, and the repair of a submarine cable is something not to be sought after. The procedure is apt to be expensive and time consuming, due to circumstances beyond control such as bad weather or lack of availability of a repair ship; and the disturbance of the cable involved in lifting it to the surface and dropping it again, possibly in something of a heap, is not desirable. It is obvious that the requirement on long life of circuit elements presents a difficult problem, especially in view of the fact that the space available for these elements is minimized in order to keep the repeater diameter small.

There was still another requirement on circuit elements, that of ruggedness. The stresses involved in laying cables in deep waters are quite considerable. The cable is under a tension of several thousand pounds and "incidents" might occur which would have no effect on an ordinary cable but which might result in dangerous shocks to the delicate elements of the repeater.

Also, as a consequence of the cable tension, the armor unlays somewhat; and this imposes twist and elongation on the interior structure, either coaxial circuit or repeater housing. The cable circuit can be designed to withstand this distortion, but the repeater housing is much more susceptible to damage from this cause.

The Repeater Housing

The requirements on flexibility and water-tightness under ocean bottom pressures were the factors of outstanding influence in the design of the re-

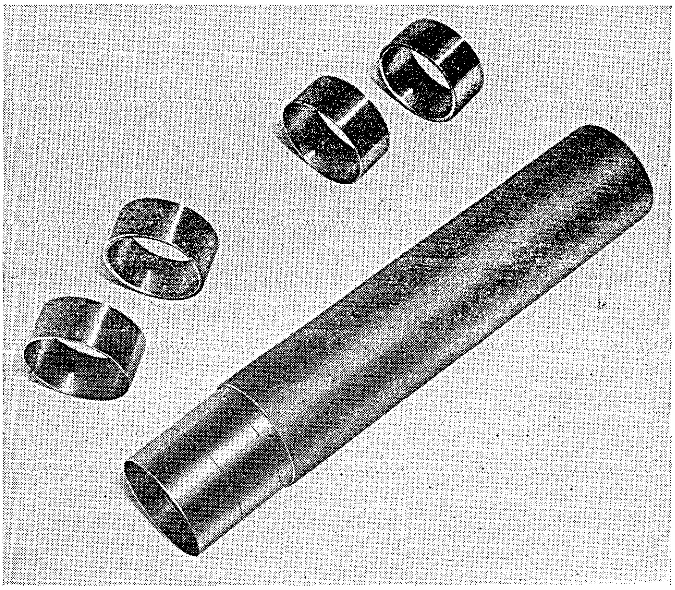


Fig. 2—Steel rings and copper envelope of repeater.

peater housing and of the end seals by means of which the cable enters and leaves the housing. In the present form the housing consists of a long tube of soft copper $1\frac{3}{4}$ inches in diameter and .03 inch thick, supported internally against collapse under sea bottom pressure by an assemblage of abutting steel rings, each $\frac{3}{4}''$ wide, and given a degree of rigidity by means of thinner steel rings of the same width overlaying and staggered relative to the thicker rings. When this structure is sealed at the ends it is capable of withstanding pressures as high as 10,000 pounds per square inch and it can be bent to a radius as small as three feet without undue distortion of the copper envelope. Details of the structure are shown in Fig. 2.

Into each end of the housing is led the insulated conductor of the cable by means of a series of seals. The inner or vapor seal is of the glass-metal type especially developed for this particular purpose and capable of withstanding considerable hydrostatic pressures. Next in line there is a seal comprising a central brass tube and an external brass member, both vulcanized to rubber, which is joined to the insulating material of the cable. These seals are co-axial in form, the outer member in each case being brazed to the copper tube of the housing or an extension thereof. Finally a closely fitting "core tube" of copper, extending over the cable insulation for a distance of about seven feet, is brazed to an extension of the copper envelope of the housing, filled with vistac and sealed at the distant end by means of a neoprene sleeve joined firmly to the core tube and to the cable insulation.

The repeater housing and core tube are provided with corrosion protective layers and a bedding for the armor wires, the bedding over the core tube being built up in the form of a taper. The armor is a continuation of the cable armor wires with additional wires interspersed on account of the larger diameter of the repeater. To prevent twisting of the container due to the unlaying of the armor wires under tension, a second layer of wires with a direction of lay opposite to that of the main armor is employed. The repeater may be armored as part of the cable or it may be armored separately, with a stub on each end, and spliced into the cable.

The components of the housings and seals, as well as the complete armored housing, have been subjected to exhaustive tests of various sorts. The rubber-brass seal, for instance, was tested for penetration of moisture vapor over long periods of time. Methods of making this seal were checked by tension tests until a uniformly high degree of adhesion was obtained. Armored housings were tested on a laboratory setup in which laying conditions could be simulated by bending the structure under tension and in motion around a six foot diameter drum.

The Repeater Circuit

The diameter of the housing had been chosen originally on the assumption that the bulge caused by the repeater should not be more than two or three times the diameter of the cable proper in order to reduce the possibility of over-riding turns on the brake drum during laying. Mechanical tests indicated that this diameter was also safe from the standpoint of deformation of the copper envelope during bending. Accordingly, it was required that the repeater structure should be restricted in cross-section so as to fit inside this tube, with as much length as would be needed.

The problem then became one of packaging the elements involved in a high gain electron tube amplifier in the restricted space available. The method finally adopted is shown in Fig. 3. The completed amplifier consists of an articulated assemblage of composite lucite cylinders, each about five inches long, successive units being held together by a spring assembly. Each lucite cylinder contains the related electrical elements of a particular part of the repeater circuit. The groups of smaller elements are mounted rigidly in a lucite form which slides into an insulating envelope consisting of two close fitting lucite shells and is held in place by end pieces of lucite. Eight copper tapes laid in axial slots between the shells and extending over several sections, where necessary, permit electrical interconnection of the various parts. A representative assemblage is shown in Fig. 4. In the case of the Key West-

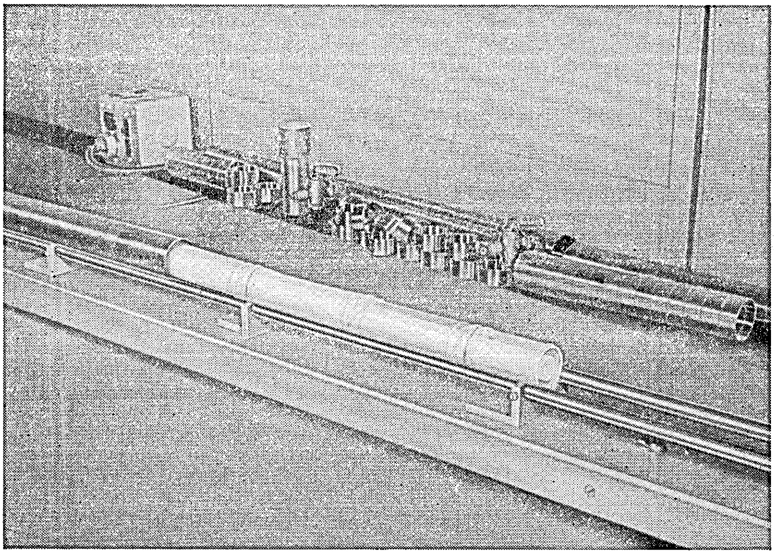


Fig. 3—View of repeater assembly.

Havana repeater the complete assemblage is eighty-four inches long and comprises fifteen sections.

Circuit Elements

Early in the development general principles were developed regarding the type of circuit best suited for underwater repeaters and on this basis requirements were established on the characteristics necessary for the circuit elements, including electron tubes, and on their arrangement in the repeater. Decisions in such matters could not be arbitrary of course, but had to be carefully worked out in order to freeze designs as early as possible so as to facilitate the start of significant life tests.

The electron tube is the most important of the elements. Work had been begun on a tube suitable for this use as far back as 1933. Thus, when a decision was made to lay the new cables, a long background of experience was drawn on in the manufacture of the tubes. Early models of the type had been operated on continuous life test for as long as ten years. Designed primarily for long life, the tube is a suppressor grid pentode with an indirectly heated cathode. Of rugged design to withstand the shocks of cable laying, the spacings between electrodes are relatively large. Unusual care was taken in manufacture to insure solid welds and to avoid the presence of loose particles. During various stages of assembly, rigorous inspections were made on all tubes by engineering personnel. Selection of tubes for use in the cable was based on a thorough examination of all details in the history of each tube,

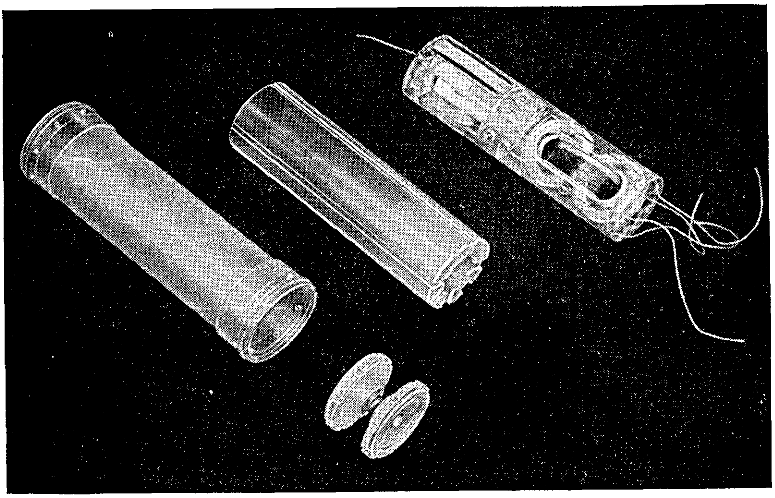


Fig. 4—Repeater network assembly.

as well as the history of the group in which it was manufactured. All tubes which were candidates for the cable were aged several thousand hours before preliminary selection was made. In addition, other tubes from the same production group were further life tested several more thousand hours to establish the quality of the group. One early decision was to power repeaters by direct current fed from land over the cable conductor. The tube heaters connected in series would furnish plate and grid potentials. This was an important factor in setting the nominal power requirements for the tube, which are about $\frac{1}{4}$ ampere at 20 volts for heater supply and plate potentials of 40 to 60 volts.

While the electron tube is usually the most vulnerable element in electrical circuits from the standpoint of life, attention must be given to other elements—condensers, resistances, and coils—especially where they are subject to long continued application of electrical potentials, as in the case of power separation filters. The factors that determine the life and performance of these elements are not completely under control. It was felt, however, that the best assurance on dependability could be obtained by careful, conservative design and by manufacturing and assembling the elements into repeaters under the best possible conditions of cleanliness. An air conditioned space was provided for this purpose at the Murray Hill Laboratory. In addition to cleaning the air in this space, precautions were taken against the entrance of dirt by other means, for example, on the clothing or persons of operators.

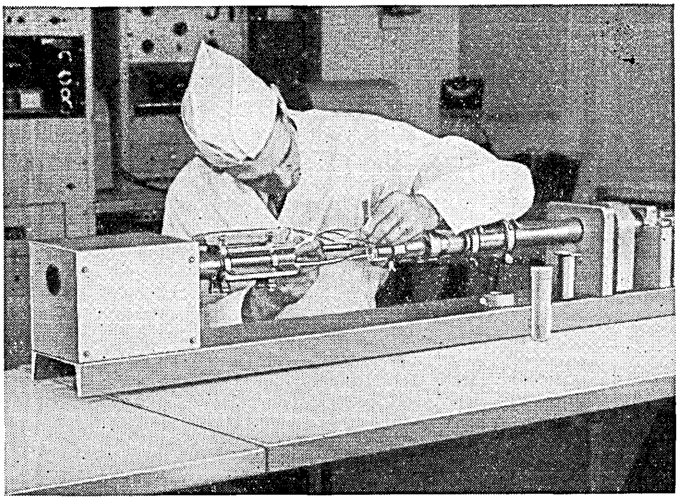


Fig. 5—Assembly of the repeater.

The humidity of the air was carefully controlled to prevent contamination from perspiration during handling of the parts. Manufacture was carried out by selected workmen, and the product was inspected at various stages by engineers. A view of one of the operations is given in Fig. 5.

Field Trials

Simulated laying tests in the laboratory covered a period of several years. Special pains had been taken to include as far as possible all aspects of the laying operation, even those which were judged to be free of hazard to the repeater. Tests were also made to determine the effect, if any, of the laying operation upon the electrical transmission characteristics of the cable. Likewise, comprehensive electrical tests had been made to insure that no unexpected effects would be encountered due to the immersion of the repeater in water.

A large scale test was needed, however, to establish the practicability of the repeatered cable. This is because of the fact that during the laying operation the suspended length of cable trailing the ship may be as great as ten nautical miles or more, and in this length there occur complex mechanical phenomena which cannot be simulated in the laboratory with a great degree of assurance. After preliminary trials from a barge in Long Island Sound, a deep water test of the repeatered cable was made in 1948 in the Bahamas. The cable ship LORD KELVIN of the Western Union Telegraph Company was chartered for the purpose. Lengths of cable up to 15 n.m. were paid out along with repeaters in depths of water up to 2 n.m. Several repeaters were laid, measured while on the ocean bottom and then hauled back to the ship, a procedure that involves much more severe treatment than a mere laying operation. The repeater shown in Fig. 1 experienced this treatment. Tests were also made with repeater housings containing specially designed accelerometers to determine the shocks resulting from possible abuse during laying. The results indicated that the repeaters as well as the cable could take the punishment with considerable margin of safety.

The Transmission System

Designing the electrical circuit of the repeater was largely a matter of getting the most out of the long life electron tube in the way of stability of repeater gain and low modulation while obtaining as much gain as the system permits. For most efficient use of tubes and to simplify the structure a unidirectional repeater design was decided upon.

The repeater employed in the Key West-Havana cables has three stages with negative feed back, the circuit being as shown in Fig. 6. The gain frequency characteristic is shown in Fig. 7. The transmission band is from 12 kc. to 120 kc. The insertion gain at 108 kc., the top frequency employed in traffic, is 65 db. The repeater gain equalizes the loss of about 36 n.m. of cable, the attenuation frequency characteristic of which is shown in Fig. 8. For comparison the characteristics of the earlier cables are also shown.

The layout of the new repeatered cable installation is shown in Fig. 9. There are two cables, one for each direction of transmission. The East, or No. 5 cable, transmitting South, is 114.55 n.m. in length. The West, or No. 6 cable, transmits north and is 124.97 n.m. in length. Each cable has three repeaters spaced approximately 36 n.m. apart. Two of the six repeaters are in a depth of .9 n.m. and two in about .35 n.m. The last repeater in each cable is located as close as possible to deep water so as to strengthen the signal before it enters shallow water and land sections of cable where static

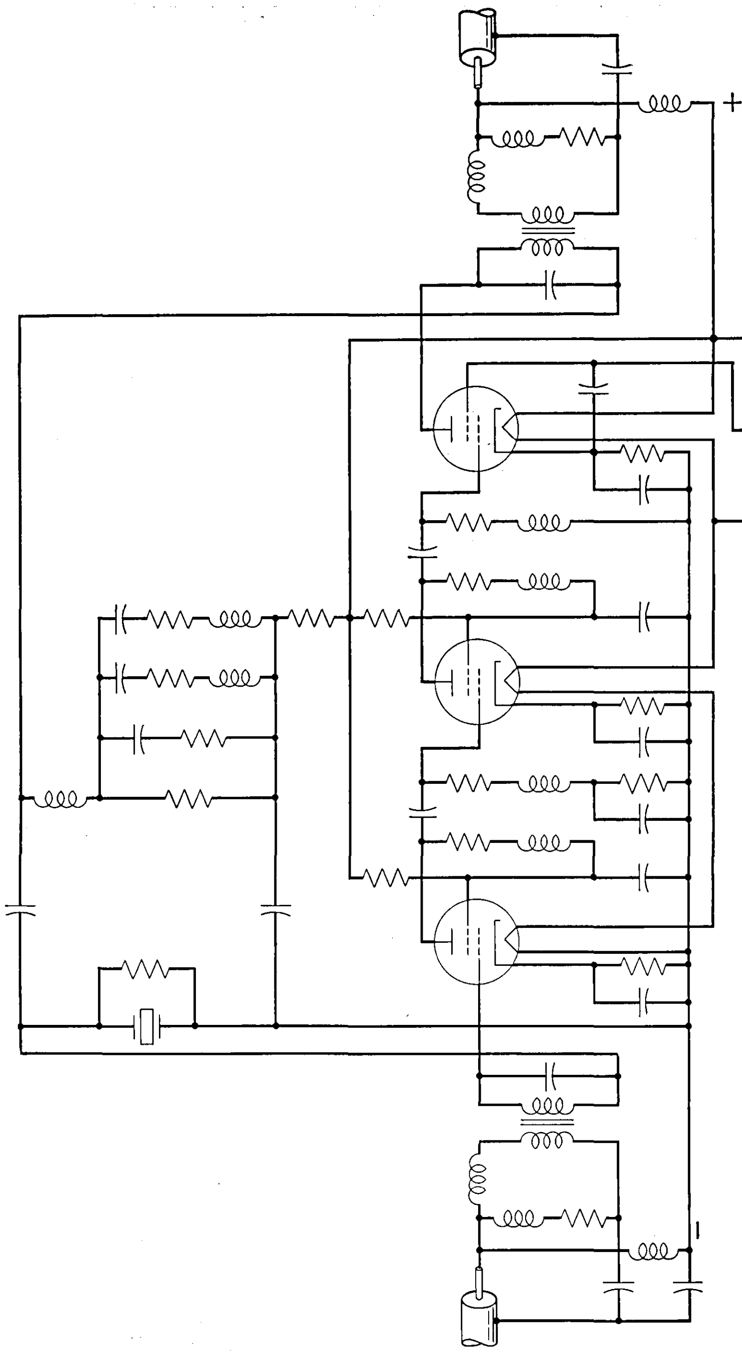


Fig. 6—Circuit of repeater.

noise and crosstalk might be picked up. In Havana the last repeater is located in a large vault on the beach. At Key West the last repeater is located close to Sand Key Light where the water begins to deepen.

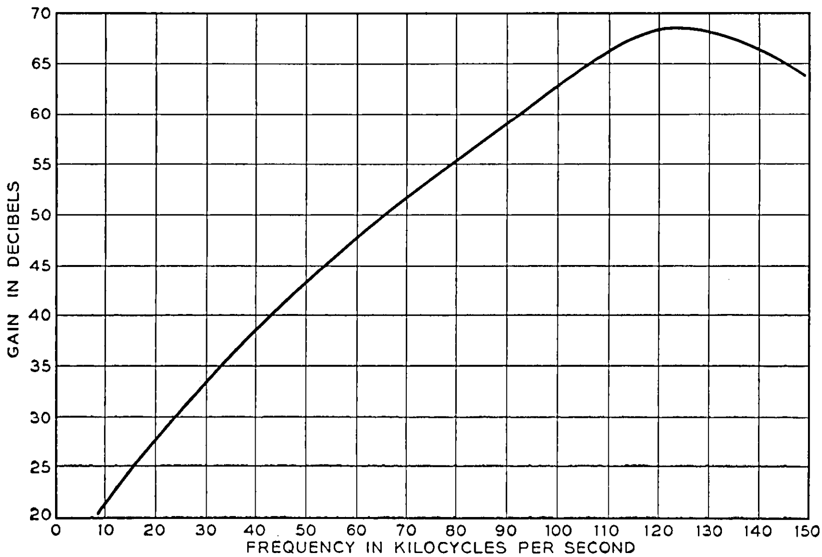


Fig. 7—Gain characteristic of repeater.

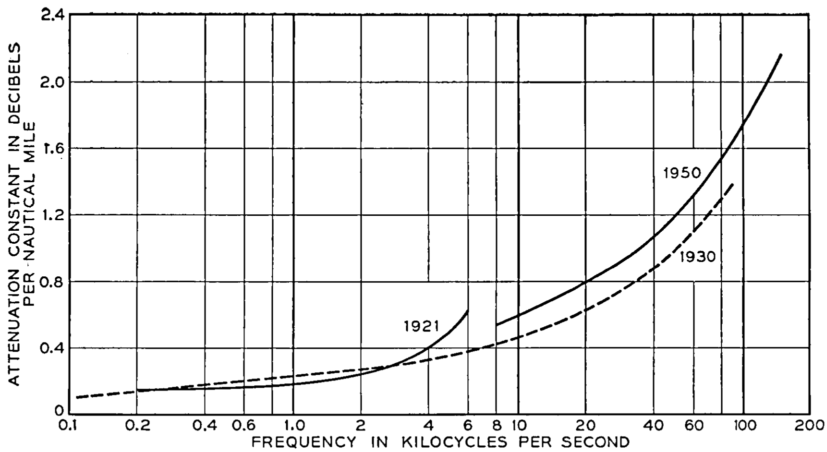


Fig. 8—Attenuation frequency characteristics—Key West-Havana Cables.

At manholes near the shore at both ends the cables are spliced directly to underground cables running in ducts to the terminal equipment at the offices, a distance of three miles at Havana and one mile at Key West. The under-

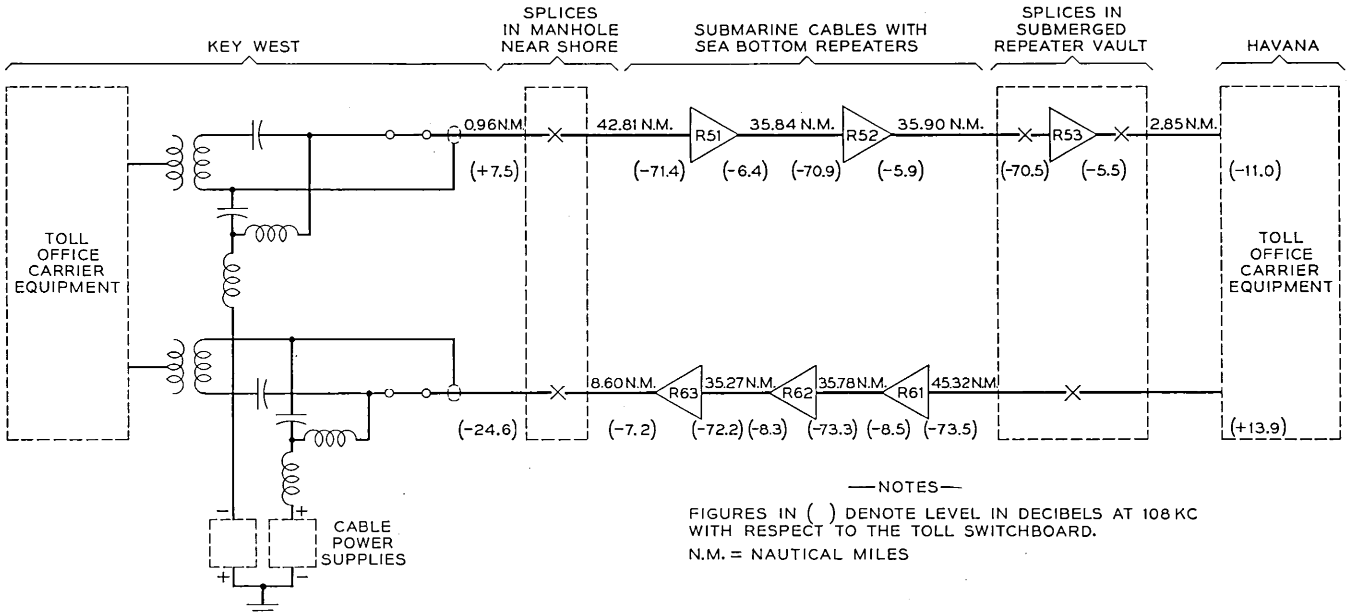


Fig. 9—Layout of system.

ground cables have the same coaxial circuit as the submarine cables but in place of the mechanical protection of jute and armor they are provided with electrical protection of helical steel tapes, layers of paper and over all a lead sheath.

The 12 kc. to 108 kc. passband yields 24 channels in each cable, each channel occupying a band of 4 kc. The signal-to-noise ratio for these channels is about the same as for the same length of high grade carrier frequency circuit on land.

The Cable

The cable has a copper return, as in the case of the earlier installations, but differs from them in being insulated with polyethylene. It also involves some new principles of design that render the cable circuit less subject to change of electrical characteristics due to laying stresses. This is a matter of considerable importance in the case of a system with submerged repeaters, since after the cable has reached the bottom it is impossible to adjust the repeater to compensate for changes in cable attenuation during laying, a matter that in ordinary cables is taken care of by adjusting the equipment on shore.

In order to avoid undesirable irregularities in transmission characteristics special precautions were taken during manufacture to obtain a higher than usual degree of uniformity of the cable impedance as seen by a repeater. Because of the wide transmission band, schemes heretofore employed for reducing the effect of the variation of impedance among the core lengths constituting the cable would have called for core lengths so short as to seriously increase the number of joints. The irregularities were therefore minimized by careful control of conductor and insulation diameters and by continuously insulating lengths of the order of 12 n.m., cutting them only as was necessary for handling, and reassembling the shorter lengths as far as possible in insulating order to assure random addition of reflections due to impedance irregularities. The success of this technique is evidenced by the impedance deviation curves shown in Fig. 10.

The structure of the cable is shown in Figs. 11 and 12. The central conductor consists of a solid wire .131 inch in diameter on which are laid three copper tape surrounds each .0145 inch thick and .148 inch wide, closely conforming to the solid wire. The interstices of the conductor are filled with polyethylene. The stranded conductor, .160 inch in diameter, is insulated with polyethylene to a diameter of .460 inch. Directly on the polyethylene insulation is laid the return conductor comprising six copper tapes, each approximately .016 inch thick by .241 inch wide, preshaped so that when in place they conform to the surface of the insulation. Both the return tapes and the tape surrounds of the central conductor have left hand lay. Over the return conductor is wound a teredo tape approximately .003 inch thick with overlap. Over all is the cutched jute, the armor and the outer jute serving, to which are applied the usual cable compounds. Four types of armor are employed in the cable for use in various depths of water or for special shore conditions.

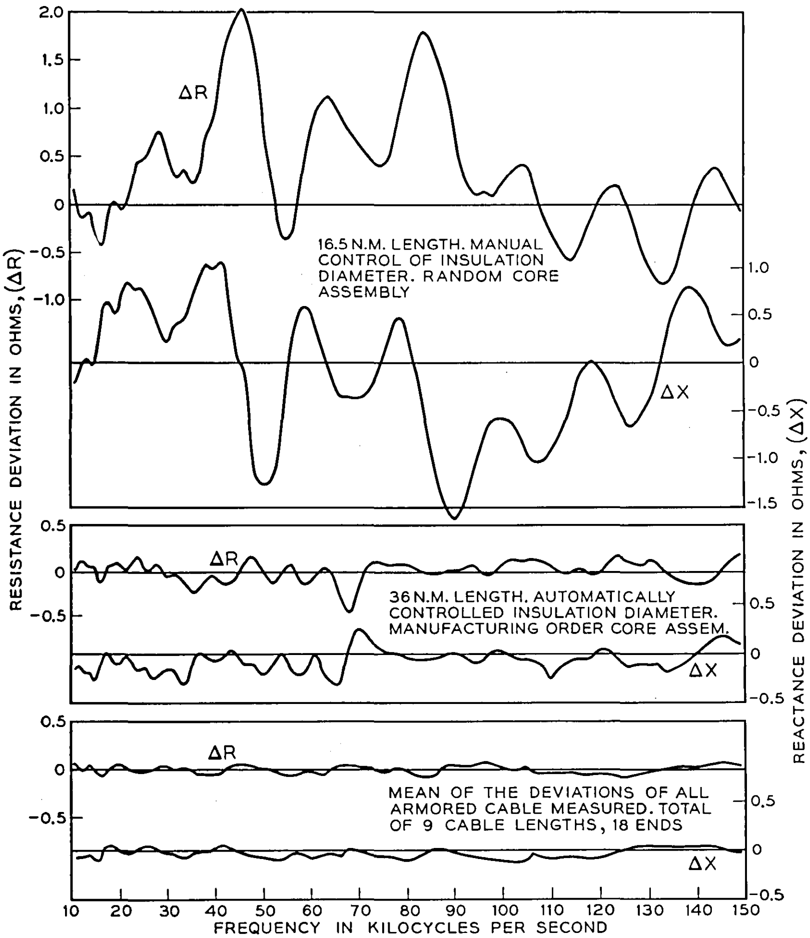


Fig. 10—Impedance deviations of cable lengths from average of all armored cable impedances.

The lengths of the various types, in nautical miles, as they appear in the two cables, starting at Key West are as follows:

Type	No. 5 Cable	No. 6 Cable
A	14.31	12.65
B	25.60	31.22
D	72.73	76.17
B	1.44	4.39
A	.16	.18
AA	.31	.36

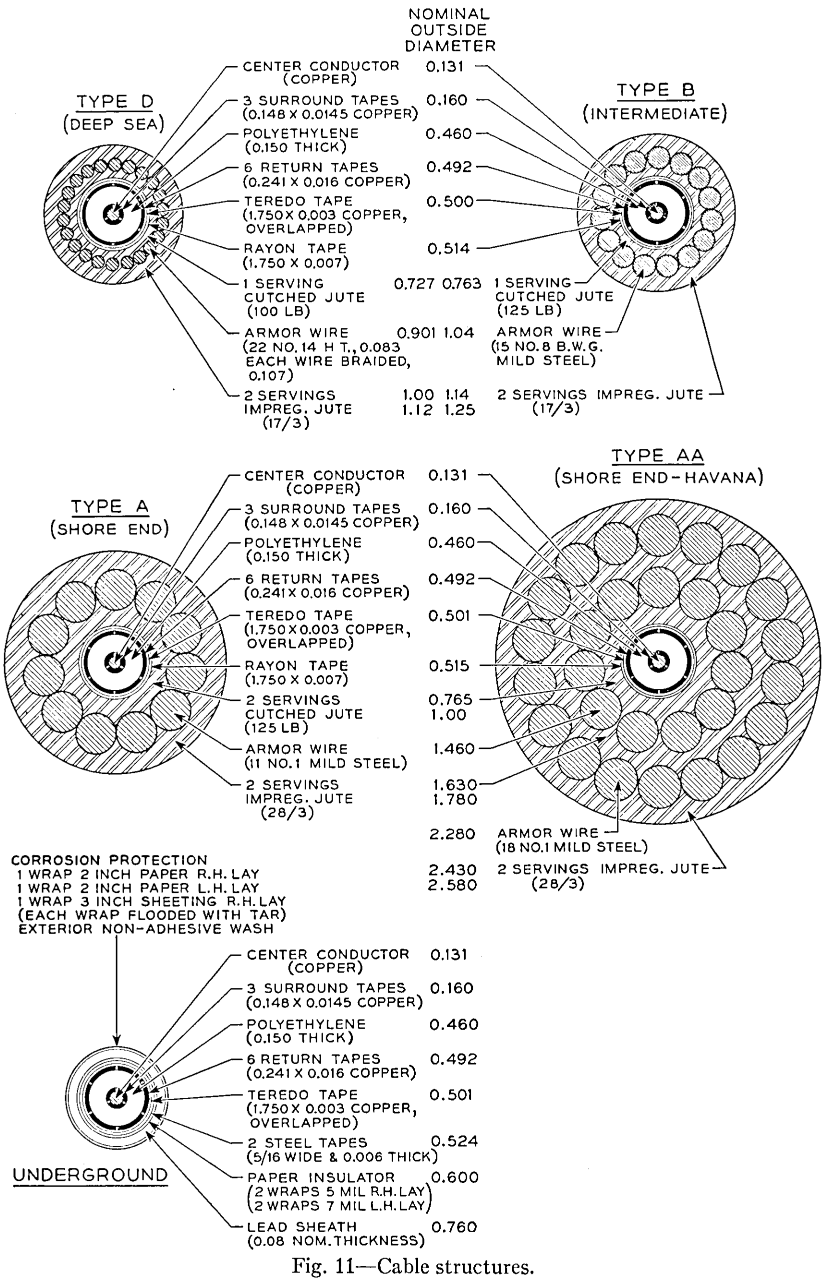


Fig. 11—Cable structures.

During the course of manufacture and in splicing on board ship, joints in the copper conductors were silver soldered. For joining the polyethylene insulation a special molding machine was designed and built by means of which polyethylene under high pressure and an elevated temperature was applied to the surfaces to be joined.

The cable was manufactured by the Simplex Wire & Cable Company of Cambridge, Mass., and incorporated the results of a cooperative development program conducted by this Company and the Bell Telephone Labora-

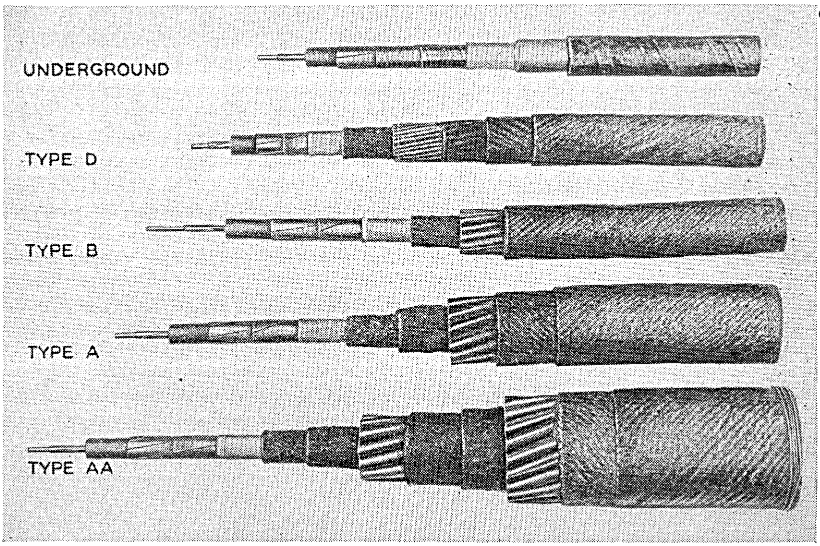


Fig. 12—Cable types.

tories. The excellent quality of the cable is a tribute to the manufacturer in this very difficult and exacting field.

Terminal Equipment

The transmission apparatus at Key West and Havana is mostly standard equipment employed in land line carrier systems, and the operations involved in combining the twenty-four voice circuits into one band and separating them again are largely conventional. Special equalizers, power separation filters and an auxiliary amplifier had to be designed and the standard transmitting amplifier used in the J system was modified to accommodate the lower frequency band. A feature of particular interest is the equipment for testing the electrical condition of the repeaters from measurements at Key West. Each repeater contains a sharply tuned circuit by means of which the gain

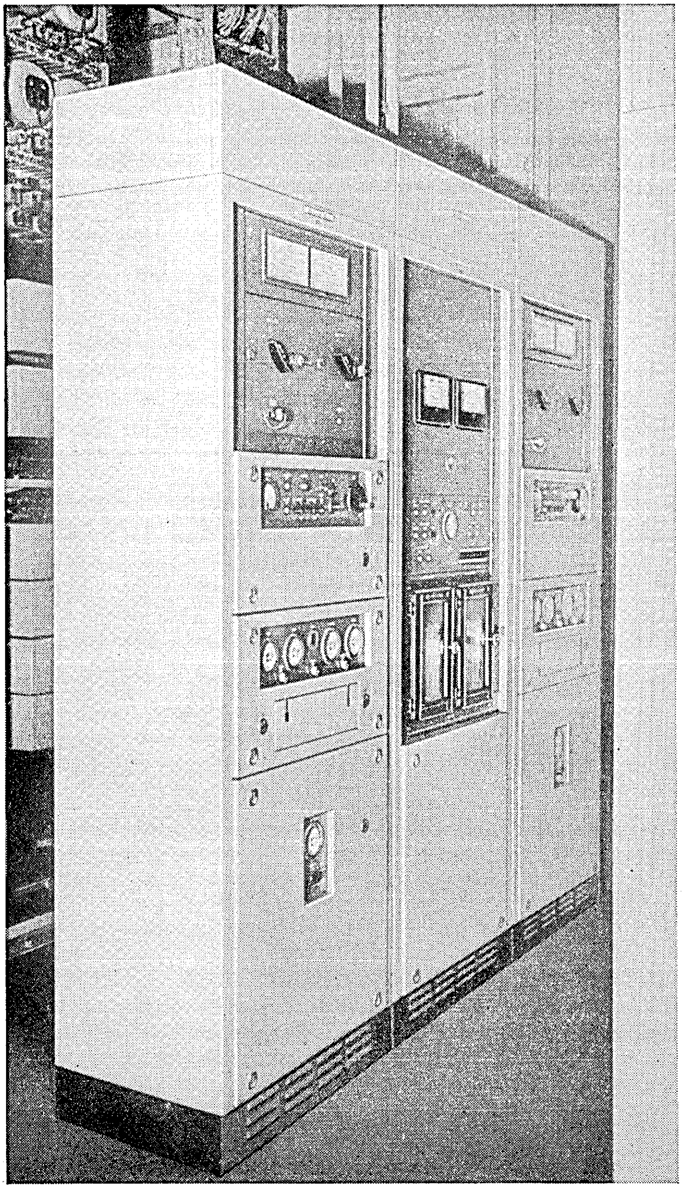


Fig. 13—Terminal power supply.

of the repeater is increased above normal at a distinctive frequency outside the transmission band of the repeater. With the aid of a loop circuit at Havana the gain with reduced feed back of the individual repeaters can be measured by scanning the test frequency region with an oscillator and detector at Key West. An indication of incipient decay of gain of any repeater is thus given.

The power for the repeaters is supplied over the cable conductor from Key West. A positive potential of about 250 volts is applied to one cable and -250 volts to the other, with a loop connection between the two cables at Havana to complete the d-c. circuit. This neutral point is also connected to ground. The current in the cable conductors is at present maintained at 0.23 ampere. A view of the rectifying and control equipment for one of the polarities is given in Fig. 13. Precautions are taken against interruption of the power supply to the cable, and sensitive controls are provided to maintain the current constant in spite of earth currents and to guard against excessive currents or potentials in the cable system in case of trouble in the power supply or in the cable itself.

Laying the Cables

The laying of the cables was completed without undue incident by the Cable Ship LORD KELVIN. The task was one of unusual difficulty since modifications had to be made in the cable laying gear, some of them untried, and it was particularly desirable that the prescribed lengths and courses be realized.

Modifications were made in the cable laying gear in order to obtain an additional margin of safety in laying repeaters. As was previously indicated, the repeater is capable of bending without harm on a diameter of approximately 72 inches; and the existing cable drum, approximately 68 inches in diameter, would have been adequate. It was felt desirable, however, to build the drum on the LORD KELVIN out to an 85-inch diameter to match the diameter of the bow sheaves. The dynamometer sheaves and the sheave leading the cable off from the brake drum presented more of a problem. The lead off sheave was replaced by a ring sheave, 85 inches in diameter, supported on wheel bearings. The frame supporting these bearings was hinged at one end and the pressure on the other end of the frame, due to the tension of the cable passing over the sheave, offered a ready means for measuring this tension. For this purpose a resistance pressure cell was employed with a recorder, which not only gave a continuous record of tension but also relayed the signals to a vertical indicator on deck for the guidance of the brake operator, and to a smaller indicator on the bridge. This arrangement is shown in Fig. 14. It is felt that it has much to recommend it over the conventional dynamometer from the standpoints of sensitivity and quickness of response.

It was important to measure the transmission characteristics of the re-

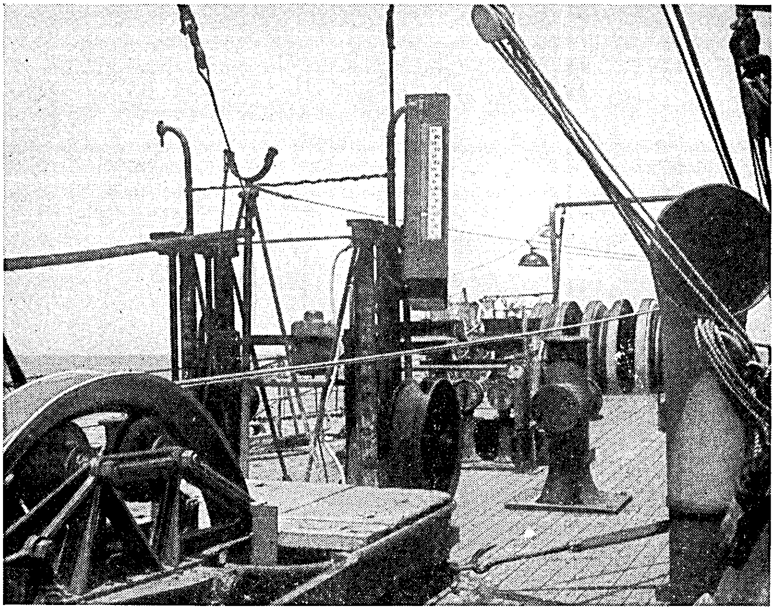


Fig. 14—Ring sheave and dynamometer scale.

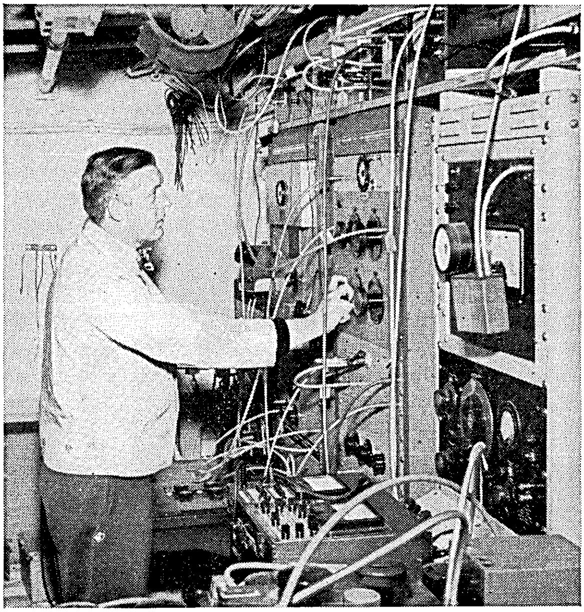


Fig. 15—Electrical laboratory on shipboard.

peatered cable before, during and after laying, and the special equipment needed for this purpose was more than could be contained in the electrician's room usually provided on cable ships. The jointers' store room was accordingly taken over and converted into an electrical laboratory, shown in Fig. 15.

The cable, loaded on board the LORD KELVIN, with the deep sea repeaters spliced in and stowed away in the tanks, arrived off Key West on April 21, 1950. Courses had been laid out for the two cables with the idea of keeping five-mile separation between the two most of the way, and five-mile separation from the nearest of the cables that constitute the rather complicated network between Key West and Havana. It is hoped thereby to avoid having the new cables picked up by mistake in connection with the repair of other cables and to avoid confusing the two cables in case either one of them is in need of repairs.

The stretch of water between Key West and Sand Key Light, a distance of about 8 n.m., is too shallow for the operation of a ship of the size of the LORD KELVIN so the sections of the two cables in this area had been laid from barges by the Long Lines Department of the American Telephone and and Telegraph Company. At Havana a new landing place had been selected. Experience with existing cables which land in Havana Harbor indicate that considerable deterioration of armor takes place in this locality and there is also the anchor menace. In addition, closeness to an existing carrier frequency cable might have given rise to undesirable crosstalk. The new landing place at the foot of B Street in Havana is about three miles from the harbor. Figure 16 shows the landing site as viewed from the cable ship during the laying operation. A view of the interior of the vault on the Havana shore is given in Fig. 17.

After putting out mark buoys at strategic points and at intervals of about 12 n.m. along the course of the cable, the Key West shore end of No. 5 cable was picked up at Sand Key and spliced on to the cable in the tanks. Then 32 n.m. of this cable were paid out, and the end buoyed at the point of final splice. The ship then proceeded to Havana, and landed the manhole repeater which was spliced to the underground cable to the office. The ship then floated the end of the cable ashore on barrels with the aid of a line operated by a winch manned by Cuban Telephone Company personnel. As soon as the end reached shore it was spliced to the repeater, the barrels were cut off, the cable dropped to the bottom and the cable on shipboard was paid out until the point of final splice was reached, where the end on board was spliced to the previously buoyed end to complete the connection between Key West and Havana.

The ship then returned to Havana, landed the end of No. 6 cable by means

Fig. 16—Havana landing site.

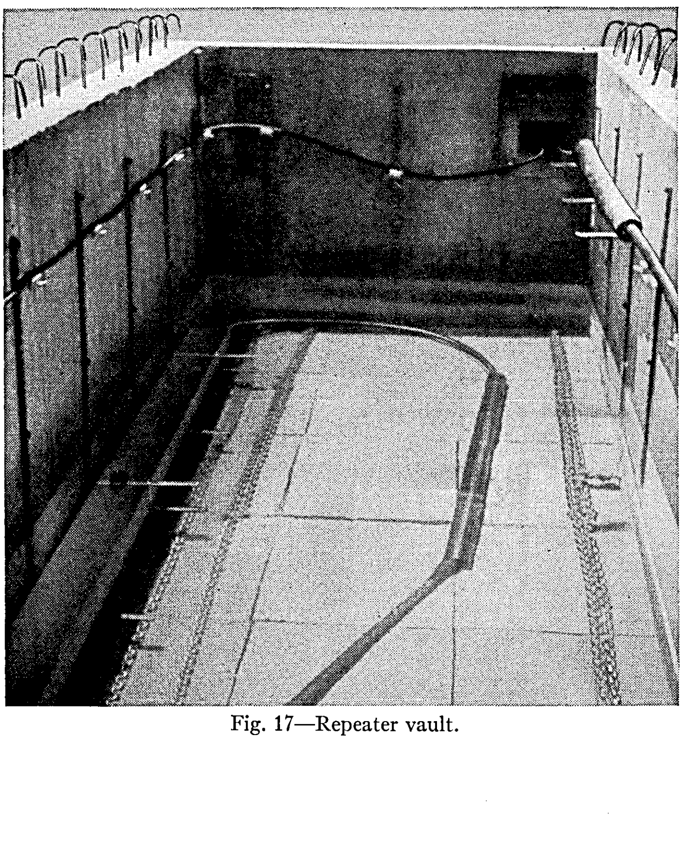


Fig. 17—Repeater vault.

of barrels and winch line and paid out cable to the point of final splice, which in this case was about four miles from Sand Key. The end of No. 6 at Sand Key was then picked up, a repeater spliced to it and to the end of the cable in the tank, and the latter was paid out to the point of final splice. Within a short time after completing this splice, insulation measurements had been completed on the two cables, the power supply was connected in to activate the repeaters, and conversation over the cable system took place.

Careful attention was given to the amount of slack paid out, that is the excess of cable length over the actual distance traversed. This latter distance is usually determined by observing the length of a taut wire paid out continuously during the laying. In the absence of taut wire gear other methods had to be devised. Observations on buoys by radar and range finder provided almost continuous information regarding the position of the ship and gave satisfactory information on slack. The conditions for cable laying between Key West and Havana are far from good. The Gulf Stream is swift and erratic. The velocity of the current at any particular point as indicated by the stream at the buoys was found to vary considerably over a fairly short period of time. As an indication of the degree of precision obtained by careful navigation of the ship, the final results show that in each of the cables the specified length was missed by only .2 n.m., which is quite an unusual achievement.

Acknowledgment is made to the Western Union Telegraph Company, the owners of the LORD KELVIN, for their cooperation in providing the special equipment for the ship and to the Captain of the LORD KELVIN, his staff and crew, for the very satisfactory performance of the laying operation.

Since the installation of the new system, it has been subjected to comprehensive tests involving measurement of noise and intermodulation between channels as well as precise measurements at numerous frequencies of net loss of the repeatered cables at intervals of time. The system has proved to be very stable and has met the requirements laid down for it. This was as expected. Nothing unfavorable to the submerged repeater has made itself felt; but, in accordance with conservative submarine cable tradition, its performance will be critically observed over a period of time.

Activity in the development of the repeatered cable and the conduct of the Key West-Havana project centered in a small group of Bell Telephone Laboratories' engineers specializing in submarine cable work and drawing on the advice and help of other groups of various backgrounds. At times, especially when troubles were encountered, the contributions of these groups were of tremendous importance and considerable in extent. The writer, as project engineer, takes this opportunity of acknowledging the assistance of the numerous individuals involved in the success of the undertaking.

Theory of the Negative Impedance Converter

By J. L. MERRILL, Jr.

(*Manuscript Received July 3, 1950*)

This paper presents a relatively new approach to the solution of negative impedance problems related to vacuum tube circuits. The approach consists of reducing the vacuum tube circuit of a device for producing negative impedance to an electrically equivalent four-terminal network from which the stability and the operation of the device as an element in a system can be predicted accurately. The theory is of interest at this time because a negative impedance repeater, the E1, has recently been developed for use in the exchange plant. It has been found that such a repeater can be placed in series with a voice frequency telephone line to provide transmission gains which are ample for many purposes.

Introduction

A NEW type of telephone repeater known as the E1 has been developed recently to meet the large demand in the exchange area for an economical means of providing transmission gains of about 10 db in two-wire telephone lines. This repeater costs less than the 22 type which has been the standard two-wire, two-way means of amplifying voice currents in the Bell System. The difference in cost is made possible by a difference in operating principle. The E1 repeater employs a type of feedback amplifier the action of which can be said to have the properties of a negative impedance converter. It is the purpose of this paper to describe the operation of the negative impedance converter, which is a device for transforming positive impedance into negative impedance.

Negative impedance like positive impedance can have two components: reactance and resistance. The reactance component can be either positive or negative. However, for an impedance to be negative the resistance component should be negative at some frequency in the range from zero to infinity.

The idea of negative resistance originated over 30 years ago, and in the beginning was associated with the concept of resistance neutralization. This concept grew from the observation that a two-terminal device could be found which had an unusual property when inserted in series with a single mesh circuit: it could produce the same flow of current as would flow otherwise at some frequency provided a resistance R were removed from this mesh. Apparently, the addition to a circuit of a two-terminal element could neutalize an amount of resistance equal to R. Thus within certain frequency limits this two-terminal device could be treated as a negative resistance equal in magnitude to $-R$.

In the early days of vacuum tube development the negative resistance effect was considered to be an important one. Possibly the regenerative vacuum tube circuits associated with the early radio receivers stimulated interest in the subject. One of the first text books on the theory of vacuum tube circuits[1] devoted about as much space to regenerative means for producing negative resistance as it devoted to the theory relating to any one of the more conventional devices—namely: amplifiers, oscillators, modulators and detectors. In spite of the interest in the subject, little practical use was made of the negative resistance theory.

Negative resistance cannot be completely disassociated from reactance. A vacuum tube circuit arranged to develop negative resistance will present reactance as well at some frequencies, and the effect on an external circuit at these frequencies will be that of taking away resistance and adding or subtracting reactance. At high and low frequencies the circuit may present a positive impedance. Consequently, the term negative impedance is used herein to designate the effect produced by a two-terminal device which has the property of negative resistance at some frequency or frequencies, negative resistance plus reactance at other frequencies and positive impedance at still other frequencies.

The Negative Impedance Converter

Heretofore, many vacuum tube circuits have been devised for converting positive impedance into negative impedance. All known simple circuits employing vacuum tubes for obtaining negative impedance, as distinguished from the combination circuits which can be made up either of vacuum tubes or of negative elements found in nature, have much in common and can be treated as the same type of arrangement. Essentially, this type of arrangement is a feedback amplifier and can be treated as such. Recently, a new method of handling these devices has been developed which has the merit of simplifying computations in many cases. This method is based upon the fact that vacuum tube circuits devised for converting positive impedance into negative impedance can be reduced to an electrically equivalent, four-terminal network consisting of a combination of positive impedance elements together with an ideal negative impedance converter.

This ideal converter, Fig. 1(a), resembles a form of transformer: it has a ratio of transformation of $-k{:}1$, can have four terminals and is capable of bilateral transmission. Assume that a positive impedance Z_N is connected to terminals 3 and 4 and $-kZ_N$ is seen at terminals 1 and 2, Fig. 1(b). Then it must follow from the theory described herein that if a positive impedance

[1] L. J. Peters—Theory of Thermionic Vacuum Tube Circuits—McGraw-Hill Book Company—1927.

Z_L is connected to terminals 1 and 2, a negative impedance $Z_L/-k$ will be seen at terminals 3 and 4, Fig. 1(c).

George Crisson stated[2] that there are two types of negative impedance which he defined as the series type and the shunt type respectively. His series type is the reversed voltage type of negative impedance $-kZ_N$ equivalent to $-V/I$; and his shunt type is the reversed current type $Z_L/-k$ equivalent to $V/-I$. This is a logical development considering that impedance Z equals V/I, and therefore negative impedance $(-Z)$ equals either $-V/I$ or $V/-I$ where V is the voltage measured across the impedance and I is current flowing through it.

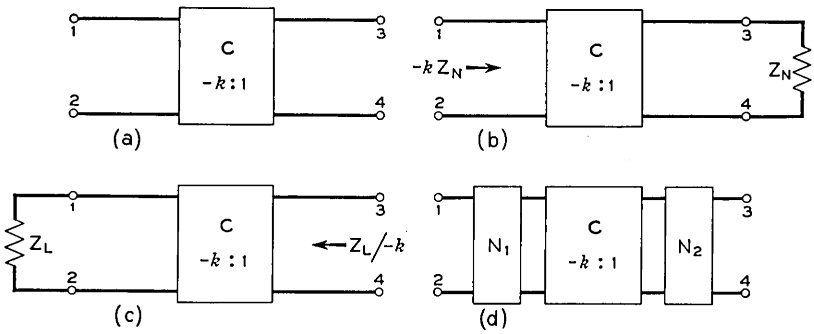


Fig. 1—The negative impedance converter.

It is important to note that k, the ratio of transformation, is of the form $A\underline{/\Theta}$ where both the magnitude A and the angle Θ are changing, however slowly, with frequency. This follows from the fact that, as will be shown, k contains a term $(\mu_1 - 1)$ where μ_1 is a voltage ratio the magnitude and phase of which can change with frequency.

The ratio of transformation k can be made to have a magnitude closely approaching unity, and at some frequency or frequencies the angle can be made zero. If k equals $1\underline{/0}$ the series type negative impedance seen at terminals 1 and 2, Fig. 1(b), will be equal to $-Z_N$, $-V/I$, where the voltage V is reversed by 180 degrees from the voltage which would appear across terminals 1 and 2 were the impedance here the positive impedance Z_N and the voltage to current ratio the conventional V/I. Likewise, the shunt type negative impedance seen at terminals 3 and 4, Fig. 1(c), would be $Z_L/-1$, $V/-I$, where here it is the current which is reversed by 180 degrees from the current which would flow through the positive impedance Z_L. Thus a strange fact is noted: multiplying a positive impedance by -1 does not yield the

[2] George Crisson—Negative Impedance and the Twin 21-Type Repeater—*B.S.T.J.*—July, 1931.

same result from a circuit viewpoint as dividing a positive impedance by -1. A positive impedance multiplied by -1 is a series type negative impedance. A positive impedance divided by -1 is a shunt type of negative impedance.

A practical (i.e., real or actual) converter circuit can be represented by Fig. 1(d). A vacuum tube circuit contains positive impedance elements. Some of these will show up in the equivalent circuit on the left-hand side of the ideal converter; others will show up on the right-hand side of the ideal converter. This will be clarified in the discussion of the E1 circuit which follows. The equivalent circuit of any practical negative impedance converter of this type can be represented by the equivalent circuit of Fig. 1(d) which shows the positive impedance elements associated with the vacuum tube in

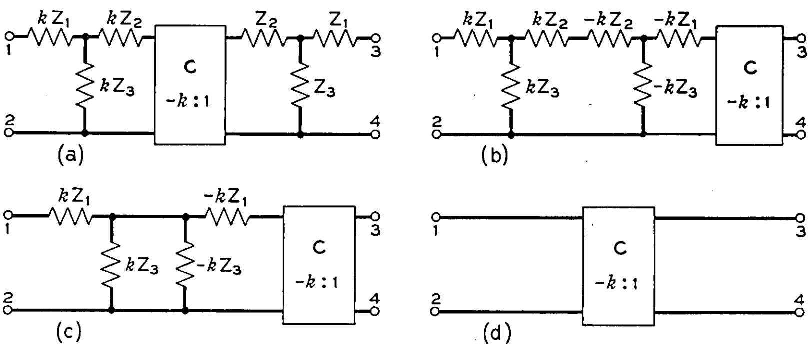


Fig. 2—Equivalent circuits.

the form of two equivalent networks (N_1 and N_2) arranged one on each side of the ideal converter (C) having a transformation ratio of $-k{:}1$.

It should be noted that should the series arms of N_1 equal k times the series arms of N_2 and should the same relationship exist for the shunt arms then the effect of these networks is cancelled, except for power dissipation, and Fig. 1(d) is equivalent to Fig. 1(a). This is illustrated in Fig. 2 where N_1 and N_2 have been represented by equivalent T networks as shown specifically in Fig. 2(a). Network N_2 can be multiplied by $-k$ and transformed to the left-hand side of the ideal converter, Fig. 2(b). The adjacent series arms of these two networks cancel each other as shown in Fig. 2(c). The shunt arms go to infinity and the other series arms also cancel leaving Fig. 2(d).

It is not possible to cancel N_1 and N_2 perfectly in a practical circuit design. But over the frequency range of interest this could be closely approximated by making all impedances shunting the ideal converter as large as possible, and by cancelling all resistances in series in N_1 by a resistance added in N_2.

As a physical concept the idea of negative impedance is difficult to visualize because this type of impedance supplies power to an external circuit rather than dissipates it. This power is supplied either by the application of a reversed voltage or a reversed current. In spite of the difficulty which may be experienced in attempting to visualize the feedback action that takes place inside a negative impedance converter, if its equivalent circuit is known then its stability can be determined readily and its operation as a device for producing negative impedance becomes obvious.

Stability

Like any amplifier whose output connects back to its input, the negative impedance converter if not properly terminated can run away with itself and oscillate. Stability can be determined by conventional feedback theory[3]; fortunately, there is a simpler criterion for determining stability. Consider again the ideal converter, Fig. 1(b). Assume that Z_L, not shown on Fig. 1(b), is an impedance connected to terminals 1 and 2. Consider the circuit mesh formed on the left-hand side of the ideal converter by the connection of Z_L to terminals 1 and 2. Here a negative impedance $(-kZ_N)$ is seen looking into terminals 1 and 2, and a positive impedance (Z_L) is seen looking away from them. The total impedance in this mesh is $Z_L - kZ_N$. If $k\,Z_N$ should equal Z_L then the total impedance would be zero; and a voltage inserted in series with this mesh would call for infinite current, a situation obviously impossible. Thus it becomes evident that kZ_N should not equal Z_L; or, what is the same thing, the ratio kZ_N/Z_L should not equal $1\underline{/0}$ if the system is to be stable. Furthermore, it can be shown that for an ideal converter the ratio kZ_N/Z_L contains the characteristics of the feedback factor $(\mu\beta)$ of the amplifier in the converter. In view of this fact, it might be expected that Nyquist's rule[4] for stability in feedback amplifiers could be paraphrased as follows: *To obtain stability in an ideal negative impedance converter the locus of the ratio kZ_N/Z_L over the frequency range from zero to infinity must not enclose the point $1\underline{/0}$.*

The same general rule for stability can be arrived at by connecting an impedance Z_N to terminals 3 and 4 of Fig. 1(c) and by considering the circuit mesh formed by $(Z_L/-k) + Z_N$. It should be noted in this case that $Z_L/-k$ calls for a flow of current 180 degrees out of phase from that which would flow through Z_L/k. This means that where the phase angle of Z_L/k equals that of Z_N the magnitude of Z_L/k must be greater than that of Z_N, which is another way of saying that at this phase angle the magnitude of kZ_N/Z_L must be less than unity.

[3] H. W. Bode—Book—Network Analysis and Feedback Amplifier Design—D. Van Nostrand Company, Inc.—1945.

[4] H. Nyquist—Regeneration Theory—*B.S.T.J.*—Jan., 1932.

From a practical engineering viewpoint there is a simple criterion for judging stability. It can be stated as follows: The ideal negative impedance converter will be unconditionally stable provided that the magnitude of kZ_N/Z_L is less than unity at any frequency where the angle of this ratio is zero.

These same conditions for stability apply to any practical (i.e., real or actual) converter circuit, Fig. 1(d). However, Z_L must be taken as the impedance seen looking into the network N_1 from the position of the ideal converter C, and Z_N must be taken as the impedance seen looking into the network N_2 from the ideal converter C. In other words, the effect of N_1 must be included in Z_L and the effect of N_2 must be included in Z_N.

Negative Impedance Converter Circuits

Negative impedance can be produced by connecting the output of an amplifier back in series or in shunt with the input in the right phase relationship. This type of circuit can be considered as a negative impedance converter similar to Fig. 1(d) where the ratio of transformation $-k$ is of the form $-(\mu_1 - 1)$, in which μ_1 represents a function of the voltage amplification of the amplifier. The disadvantage of a transformation ratio of this kind is that it changes markedly with variations in tube constants and battery supply voltage. Such circuits present a stability problem. One solution to this problem has been described by E. L. Ginzton[5]. He reduced variations in the amplifier gain by stabilizing the amplifier itself with negative feedback and thus reduced variation in μ_1. Note that if μ_1 is set equal to 2, then $-k$ becomes equal to -1.

There is another method of using negative feedback to stabilize a circuit for producing negative impedance. This method was used in the E1 circuit and will be described in detail in connection with it. Essentially, here negative feedback is arranged together with positive feedback to produce a transformation ratio for the ideal converter of the form $-(\mu_1 - 1)/(\mu_2 + 1)$. The symbol μ_2 represents a voltage ratio. Furthermore, μ_1 equals $\beta_1\mu_2$. If β approximates unity and if μ_2 is very much larger than one, $-k$ approaches -1 in value and is relatively independent of variations in tube constants and battery supply voltage.

In order to illustrate how the equivalent circuit of a negative impedance converter can be derived, consider a circuit credited to Marius Latour about the year 1920, Fig. 3(a). Figure 3(b) is a schematic representation of Fig. 3(a) in the manner originated by G. Crisson. With reference to Fig. 3(b), the polarity of amplifier A is assumed such that, at the instant the a-c current I_1 flows in the direction indicated, the current I_2 will flow in the direction

[5] E. L. Ginzton—Stabilized Negative Impedances—*Electronics*—July, 1945.

shown and the voltages V_2 and $\mu I_1 Z_N$ will have the polarities shown. For the moment it will be assumed that the input impedance of amplifier A is infinitely high, an assumption which will be modified later to agree with actual circuits. The symbols Z_N and Z_1 represent impedances, R_p represents the plate resistance of the output tube (T_2) and $\mu I_1 Z_N$ represents the voltage produced in the plate circuit of the output tube of amplifier A, the tube

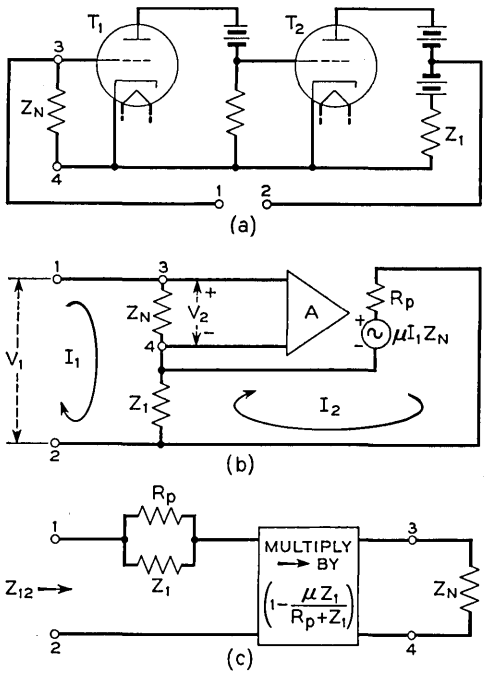


Fig. 3—Latour's circuit.

T_2 of Fig. 3(a), by the voltage drop $I_1 Z_N$ across the input, grid to cathode of T_1, Fig. 3(a). The mesh equations for Fig. 3(b) can be written as follows:

$$\begin{aligned} V_1 &= (Z_N + Z_1)\, I_1 \qquad\qquad -Z_1 I_2 \\ 0 &= -(Z_1 + \mu Z_N)\, I_1 + (R_p + Z_1)\, I_2 \end{aligned}$$

The current I_1 becomes:

$$I_1 = \frac{V_1(R_p + Z_1)}{(Z_N + Z_1)(R_p + Z_1) - (Z_1 + \mu Z_N)Z_1} \qquad \text{Eq. (1)}$$

The impedance seen looking into terminals 1 and 2 can be written:

$$Z_{12} = \frac{V_1}{I_1} = \frac{Z_1 R_p}{Z_1 + R_p} + Z_N\left(1 - \frac{\mu Z_1}{R_p + Z_1}\right) \qquad \text{Eq. (2)}$$

Thus the impedance Z_{12} equals the impedance Z_1 in parallel with R_p which combination is in series with impedance Z_N multiplied by $1 - [\mu Z_1/(R_p + Z_1)]$. An equivalent circuit for Z_{12} is illustrated in Fig. 3(c).

Next, assume that Z_L, not shown on Fig. 3(b), is connected to terminals 1 and 2 of Fig. 3(b) and that Z_N is removed from across terminals 3 and 4. The mesh equations can be written as follows:

$$\begin{aligned} -V_2 &= (Z_1 + Z_L)\, I_1 \quad - Z_1 I_2 \\ \mu V_2 &= -Z_1 I_1 \quad + (R_p + Z_1) I_2 \end{aligned}$$

The current I_1 can be written:

$$I_1 = \frac{-V_2\,[R_p + Z_1 - \mu Z_1]}{(Z_1 + Z_L)(R_p + Z_1) - Z_1^2} \qquad \text{Eq. (3)}$$

The impedance looking into terminals 3 and 4, Fig. (3b), with the changes listed above is:

$$Z_{34} = \frac{V_2}{-I_1} = \frac{Z_L + \dfrac{R_p Z_1}{R_p + Z_1}}{\left(1 - \dfrac{\mu Z_1}{R_p + Z_1}\right)} \qquad \text{Eq. (4)}$$

Hence, if the circuit of Fig. 3(b) is redrawn as a four-terminal network as shown in Fig. 4(a) with Z_2 added to represent the input impedance of amplifier A the equivalent circuit of this network can be represented by Fig. 4(b). The equivalent circuit consists of two positive impedance networks, one on each side of an ideal converter. The ratio of transformation of this ideal converter is of the form $-(\mu_1 - 1){:}1$. Looking into terminals 1 and 2, Fig. 4(b), a series type negative impedance will be seen and looking into terminals 3 and 4 a shunt type negative impedance will be seen. The proof that these impedances are negative and of the reversed voltage or reversed current type has been established by H. W. Dudley, F. H. Graham and R. C. Mathes for similar circuits and will not be taken up here, although the fact could be derived simply from equations (1), (2), (3) and (4). The purpose of this discussion is to illustrate the simplicity with which an equivalent circuit can be derived, and to point out the value of the concept of the ideal negative impedance converter.

A well known circuit which can be used as a negative impedance converter is the circuit of the 21 type repeater. In this circuit the output of the amplifier is connected back to the input through a bridge type of arrangement referred to as a hybrid coil. Negative feedback and positive feedback can be developed across this coil between the amplifier output and the amplifier input. This device was used as a negative impedance converter by Crisson in his twin 21-type repeater.

There is another type of negative impedance converter circuit which should not be confused with the converter circuits mentioned heretofore. This type was disclosed by Charles Bartlett in the March 1927 issue of the

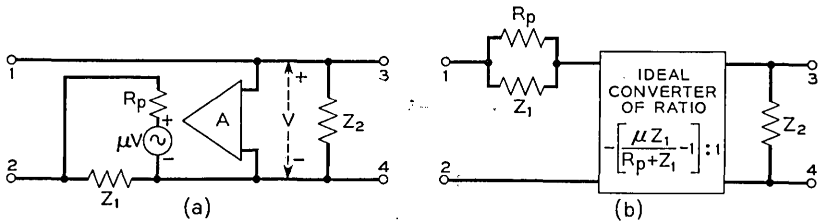


Fig. 4—A negative impedance converter and its equivalent circuit.

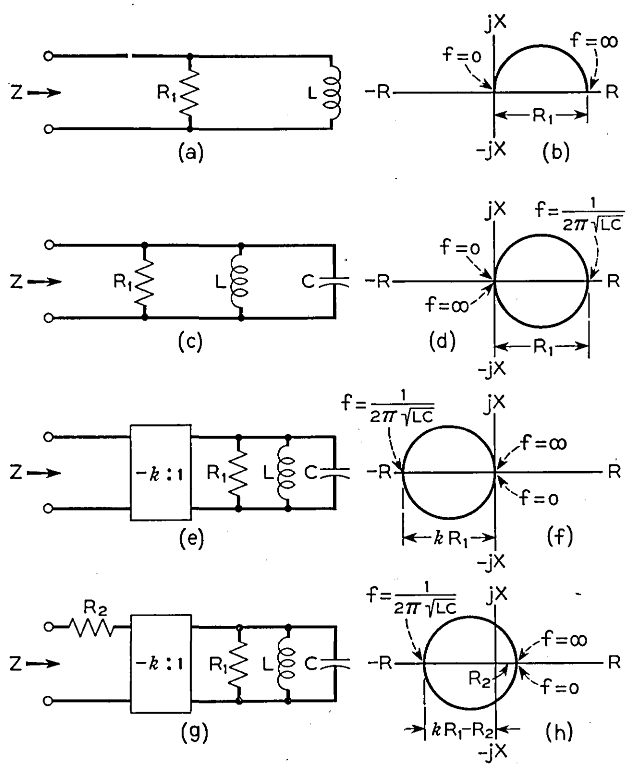


Fig. 5—Impedance loci.

Journal I.E.E. pages 373 to 376. The Bartlett converter consists of a T network of equal resistances, one of them negative. To construct such a converter there must be available a negative resistance element. Over a finite frequency band this negative resistance can be produced by a converter circuit such as that of Fig. 4. However, in this case Bartlett's circuit becomes in a sense, a converter within a converter.

The Negative Impedance Locus

Before the E1 type of converter is described, it would be well to consider, in general, the impedance characteristic which can be produced by a negative impedance converter. The shape of the impedance characteristic over the frequency range, zero to infinity, looking into terminals 1 and 2 or into terminals 3 and 4 of a negative impedance converter will be called the negative impedance locus. It is convenient to plot this locus in the polar form with frequency as a "running" parameter. The locus can be derived for any circuit by means of the theory outlined by K. G. Van Wynen for positive two-terminal impedances.[6]

For example, consider the converter of Fig. 4. Assume that an impedance such as that shown in Fig. 5(a) is connected to terminals 3 and 4 of Fig. 4; assume Z_2 of Fig. 4 is a capacitance and that both R_p and Z_1 are resistances. Now Fig. 5(a) represents a two-terminal network made up of a resistance shunted by an inductance. The locus of this positive impedance plotted on the R and jX plane over the frequency range from zero to infinity is shown in Fig. 5(b). At zero frequency the impedance is zero; at infinite frequency the impedance is R_1. If to the network of Fig. 5(a) a capacitance C is added, to represent Z_2 of Fig. 4, the impedance of the network so formed, Fig. 5(c), follows the circle of Fig. 5(d). At zero frequency the impedance is zero, at the resonant frequency the impedance is R_1, and at infinite frequency the impedance is zero again. If the impedance of this network, Fig. 5(c), were viewed through an ideal negative impedance converter, Fig. 5(e), having a ratio of transformation of $-k{:}1$ where, for the moment, k is assumed to be a numeric over the entire frequency range, the impedance locus can be represented by Fig. 5(f). Of course, k will always have an angle at high and low frequencies but, to a first approximation, at least, it can be assumed that over most of the frequency range shown in the impedance diagrams of Fig. 5(f) and Fig. 5(h) k approaches a numeric. If the circuit configuration of Fig. 5(g) is created by adding resistance R_2 in series with Fig. 5(e) the impedance locus looks like that of Fig. 5(h). This is a series type of negative impedance and simulates the impedance seen over a large portion of the frequency range looking into terminals 1 and 2 of Fig. 4 with the two-terminal network of Fig. 5(a) connected to terminals 3 and 4.

[6] Design of Two-Terminal Balancing Networks—K. G. Van Wynen—*B.S.T.J.*—Oct., 1943.

By a similar analysis, the locus of the shunt type of negative impedance can be derived. With the proper choice of an impedance connected to terminals 1 and 2 of the negative impedance converter the locus of the impedance seen at terminals 3 and 4 can be made to follow a circle in the clockwise direction (at least as long as k approximates a numeric). Thus, over a portion of the frequency range the series and shunt type of negative impedance can be made to have very similar impedance characteristics. At very high or low frequencies where k has an appreciable angle there will be a distinct difference between the locus of the series and the corresponding shunt type of negative impedance. This would be expected because in one case the network impedance is multiplied by $-k$; and in the other it is divided by $-k$.

The E1 Converter

The circuit of the negative impedance converter used in the E1 telephone repeater is shown in Fig. 6(a). It consists of a transformer, two triode tubes, an RC network and an inductor. The transformer T couples the cathodes of the two tubes to terminals 1 and 2. The tubes while apparently in push-pull are biased for Class A operation. The RC network couples the plate of each tube to the grid of the other. The inductor L supplies plate current.

The equivalent circuit of Fig. 6(a) is shown in Fig. 6(b). In obtaining this equivalent circuit the two tubes have been assumed to be identical. Thus the converter used with the E1 repeater can be reduced to a four-terminal network consisting of (reading from left to right): the equivalent circuit of the line transformer T; the two biasing resistances R_2; the two plate resistances R_p divided by $(1 + \mu_2)$; the ideal negative impedance converter C of ratio $-(\mu_1 - 1)/(\mu_2 + 1)$ to 1; the elements of the RC coupling arrangement which appear in shunt across terminals 3 and 4; the inductor L also shunted across these terminals; and the capacitor C_X which has been added to represent both the distributed capacitance of the windings of L and the capacitance between vacuum tube plates.

It should be noted that μ_2 is the amplification factor of each tube; and that μ_1 equals $\beta_1\mu_2$ where β_1 is a proportionality factor representing the fraction of the voltage, between the plate of one tube and ground, which is fed back to the grid of the other tube. The value of β_1 depends upon the values of C_1, R_3, R_5 and C_2 of the RC coupling circuit. If β_1 approaches unity in value then μ_1 approximates μ_2. If this is so and if both μ_1 and μ_2 are relatively large in magnitude compared to unity then the ratio of transformation, $-(\mu_1 - 1)/(\mu_2 + 1)$ to 1, approaches although it cannot equal $-1:1$.

As an illustration of how the elements in the E1 circuit may be proportioned, consider the practical design of the E1 converter. Here the ratio $-(\mu_1 - 1)/(\mu_2 + 1)$ to 1 is $-0.9{:}1$ over most of the voice frequency range. This is not the over-all ratio of transformation of the device, but only the ratio of the ideal converter C, Fig. 6(b). The ratio of transformer T and the effect of the other circuit elements must be considered in determining the over-all effect of the converter from terminals 1 and 2 to terminals 3 and 4. The transformer ratio is 1:9 from terminals 1 and 2 to the tube

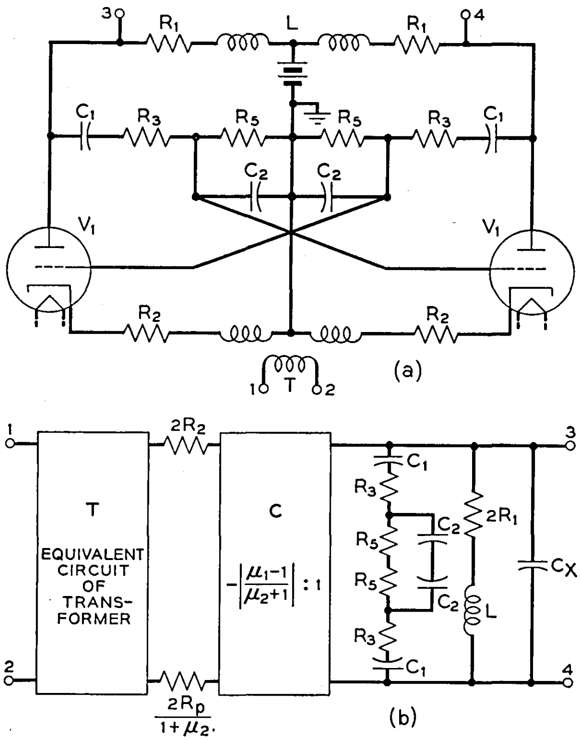


Fig. 6—The E1 converter circuit.

cathodes. The shunt arms of the networks on both sides of the ideal converter C are relatively high compared to the impedances between which this converter has been designed to operate at voice frequencies. Therefore, these shunt arms can be disregarded at voice frequencies although at frequencies above and below the voice band these shunt arms represent a problem for the circuit designer from the viewpoint of stability. In the actual E1 circuit the series arms such as $2R_p/(1 + \mu_2)$ and $2R_2$ could be cancelled out by adding in series on the right-hand side of the ideal converter a resistance of 1800 ohms (not shown in Fig. 6). The final result is that the impedance seen, looking into terminals 1 and 2 of the E1 converter when 1800 ohms plus a network Z_N is connected to terminals 3 and 4, equals $-0.1Z_N$ within a reasonable percentage of error over the frequency range from about 300 to 3500 c.p.s. for values of negative impedance from about 100 to 2000 ohms.

In the practical design two line windings instead of the one shown connected between terminals 1 and 2 in Fig. 6(a) are provided on transformer T. In practice one of these windings is inserted in each side of the telephone line in a balanced arrangement. Terminals 1 and 2 are thus effectively connected in series with the line, and the E1 repeater presents to the telephone line a reversed voltage type of negative impedance $(-V/I)$, which is the means of introducing additional power in the line thereby providing a transmission gain. The value of the negative impedance is controlled by a

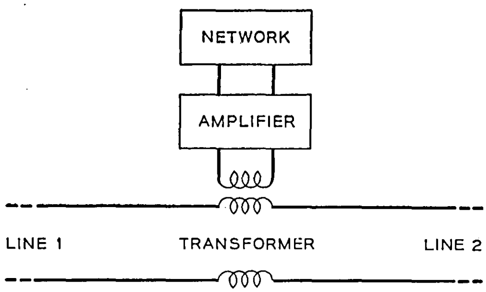


Fig. 7—E1 telephone repeater.

network connected to terminals 3 and 4. Thus the E1 repeater consists of a transformer, an amplifier unit and a gain adjusting network (Fig. 7). The transformer, the amplifier unit and part of the network make up the negative impedance converter under discussion.

For overload conditions the action of a repeater of this kind differs markedly from what might be expected from a conventional amplifier. What can be expected of a conventional amplifier is well known. An idea as to the performance of a negative impedance converter under overload conditions can be had from the following example: Assume that terminals 1 and 2 of the converter are connected in series with a telephone line and a network is connected to terminals 3 and 4 so that a reversed voltage type of negative impedance of a value less than the line impedance is inserted in the line. The combination described will be stable, and some transmission gain will be provided by this negative impedance. If now the volume of speech on the line is increased beyond the overload point the result will be a noticeable reduction in the amount of negative impedance, and a consequent reduction in gain. Under excessive overload conditions the negative impedance becomes small; and the effect of the converter is scarcely discernible on the transmission of speech as far as gain is concerned. The harmonic distortion introduced on the line by the vacuum tubes overloading will increase to a maximum and then decrease with further increase in volume. The constants of the circuit coupling the plate to the grid of the vacuum tube or tubes in the converter will determine the maximum amount of distortion. A push-pull circuit is better from this standpoint than a single sided one. In the E1 repeater harmonics are not particularly objectionable under any condition of overload.

The E1 repeater employs a single 407A vacuum tube which is a twin triode of the 9-pin miniature type. When operated on a plate voltage of 130 volts this repeater will pass speech volumes of +10 *vu* before compression begins because of overloading in the vacuum tube.

Development of the E1 Equivalent Circuit

One purpose of this paper is to prove that Fig. 6(b) is equivalent to Fig. 6(a)

Assume that an impedance Z_N is connected across terminals 3 and 4. Assume, furthermore, that the elements C_1, C_2, R_3 and R_5, which also connect across terminals 3 and 4, are all included in impedance Z_N. Assume an impedance Z_L is connected to terminals 1 and 2. The circuit of Fig. 6(a) then can be represented by Fig. 8 where the vacuum tubes have been replaced by their plate resistances (R_p) and their equivalent circuit voltages $\mu_1 e_1$ and $\mu_2 e_2$. The voltage $\mu_1 e_1$ is that voltage which appears in the plate circuit of each tube as a result of the action on the grid of all voltages between the tube plates and ground. The voltage $\mu_2 e_2$ is that which appears in the plate circuit of each tube as a result of the voltages between the cathode and ground, Fig. 6(a). The resistors R_1, R_2 and R_6 designate the resistance in the various coil windings plus other circuit resistance which might be inserted at the points indicated. The reactances X_1, X_2 and X_3 represent the effect of the self-inductance in the coil windings. The reactances M_1, M_2 and M_3 represent the effect of the mutual inductances between coil windings. The numbers on a coil winding determine the polarity of the winding with respect to other windings on the same core. As mentioned previously the vacuum tubes are operated Class A.

The basic circuit equations can be written as follows for Fig. 8:

$$\begin{aligned}
0 &= PI_1 + QI_2 - M_3(1 + \mu_2)I_3 - S(1 - \mu_1)I_4 \\
0 &= QI_1 + PI_2 - M_3(1 + \mu_2)I_3 - S(1 - \mu_1)I_4 \\
E_3 &= -M_3I_1 - M_3I_2 + (Z_L + R_6 + X_3)I_3 + 0 \\
E_4 &= -SI_1 - SI_2 + 0 + (Z_N + 2S)I_4
\end{aligned}$$

Where

$$
\begin{aligned}
P &= R_0 + R_1 + X_1(1 - \mu_1 k_1) + X_2(1 + \mu_2) \\
Q &= -\mu_1 R_1 + X_1(k_1 - \mu_1) + X_2(1 + \mu_2)k_2 \\
S &= R_1 + X_1(1 + k_1) \\
R_0 &= R_p + R_2(1 + \mu_2) \\
k_1 &= \text{coupling factor} = M_1/X_1 \\
k_2 &= \text{"} \quad \text{"} = M_2/X_2 \\
k_3 &= \text{"} \quad \text{"} = M_3/\sqrt{X_2X_3} \\
E_3 \text{ and } E_4 &= \text{applied voltages}
\end{aligned}
$$

While the derivation of most of the coefficients in these mesh equations is obvious, the derivation of P and Q may require further explanation. Coefficient P can be considered as follows:

1—The term R_0 equals $R_p + R_2(1 + \mu_2)$. The plate resistance R_p is in the plate circuit and hence stands alone. The resistance R_2 being between cathode and ground, Fig. 6(a), produces negative feedback and must be multiplied by $(1 + \mu_2)$.

2—The resistance R_1 is in the plate circuit, and here as in the case of R_p the flow of current in R_1 does not produce a voltage across the grid of the same tube through which this current flows.

3—The reactance X_1 is in the plate circuit of each tube, and by means of the mutual reactance (M_1) is coupled to the respective grid of the other tube. This coupling provides positive feedback for current flowing through X_1 which can be expressed as $-\mu_1 M_1$ or $-\mu_1 k_1 X_1$. Thus the term $X_1(1 - \mu_1 k_1)$ is derived.

4—The reactance X_2 being between cathode and ground, Fig. 6(a), provides negative feedback. Hence X_2 must be multiplied by $(1 + \mu_2)$.

Coefficient Q can be explained in similar fashion:

1—Although the flow of current through R_1 does not produce a voltage across the grid of the same tube through which this current flows, a voltage drop is produced across the grid of the other tube because of the cross coupling of these grids, Fig. 6(a). Thus voltage drop in one plate circuit appears between grid and ground of the other tube in the direction to aid the flow of current in this other mesh; hence the term $-\mu_1 R_1$.

2—Likewise, the reactance X_1 acts in the same manner as R_1 in aiding the flow of current in the other tube circuit. Furthermore, X_1 is coupled by the mutual reactance to this other mesh. These effects can be expressed as $X_1(k_1 - \mu_1)$.

3—The reactance X_2 is coupled by mutual reactance to both tube circuits. It appears in each circuit between cathode and ground in the polarity to produce negative feedback equal to $X_2(1 + \mu_2)k_2$.

In order to establish the fact that Fig. 6(a) can be represented by the equivalent circuit of Fig. 6(b) the basic mesh equations will be developed in the following manner:

First—The impedance seen looking into the converter from terminals 1 and 2, Z_{12}, will be found.

Second—An equivalent circuit which will provide this impedance will be drawn.

Third—The impedance seen looking into the converter from terminals 3 and 4, Z_{34}, will be obtained.

Fourth—The equivalent circuit for Fig. 6(a) should be the logical result.

$$Z_{12} = \frac{E_3}{I_3} - Z_L \qquad \text{Eq. (5)}$$

$$I_3 = \frac{E_3 \begin{vmatrix} P & Q & -S(1-\mu_1) \\ Q & P & -S(1-\mu_1) \\ -S & -S & (Z_N + 2S) \end{vmatrix}}{\Delta} \qquad \text{Eq. (6)}$$

where:

$$\Delta = -M_3 \begin{vmatrix} Q & -M_3(1+\mu_2) & -S(1-\mu_1) \\ P & -M_3(1+\mu_2) & -S(1-\mu_1) \\ -S & 0 & Z_N + 2S \end{vmatrix} + M_3 \begin{vmatrix} P & -M_3(1+\mu_2) & -S(1-\mu_1) \\ Q & -M_3(1+\mu_2) & -S(1-\mu_1) \\ -S & 0 & Z_N + 2S \end{vmatrix} + (Z_L + R_6 + X_3) \begin{vmatrix} P & Q & -S(1-\mu_1) \\ Q & P & -S(1-\mu_1) \\ -S & -S & Z_N + 2S \end{vmatrix}$$

Solving for I_3

$$I_3 = \frac{E_3(P-Q)\left[(P+Q)(Z_N+2S) - 2(1-\mu_1)S^2\right]}{(P-Q)\left[[Z_L + R_6 + X_3][(P+Q)(Z_N+2S) - 2S^2(1-\mu_1)] - 2M_3^2(1+\mu_2)(Z_N+2S)\right]} \qquad \text{Eq. (7)}$$

$$\frac{E_3}{I_3} = (Z_L + R_6 + X_3) - \frac{2M_3^2(1+\mu_2)(Z_N+2S)}{(P+Q)(Z_N+2S) - 2(1-\mu_1)S^2} \qquad \text{Eq. (8)}$$

$$Z_{12} = R_6 + X_3 - \frac{2M_3^2(1+\mu_2)(Z_N+2S)}{(P+Q)(Z_N+2S) - 2(1-\mu_1)S^2} \qquad \text{Eq. (9)}$$

where

$$P + Q = R_0 + (1 - \mu_1)S + X_2(1 + \mu_2)(1 + k_2) \quad \text{Eq. (10)}$$

Substituting for $P + Q$ and rearranging

$$Z_{12} = R_6 + X_3 \left| \frac{R_0 + X_2(1 + \mu_2)(1 + k_2 - 2k_3^2) + \dfrac{(1 - \mu_1)SZ_N}{Z_N + 2S}}{R_0 + X_2(1 + \mu_2)(1 + k_2) + \dfrac{(1 - \mu_1)SZ_N}{Z_N + 2S}} \right| \quad \text{Eq. (11)}$$

Equation (11) can be rewritten in a form from which the equivalent circuit can readily be constructed. This form is given below in Equation (12)

$$Z_{12} = R_6 + \left| \frac{1}{\dfrac{1}{X_3} + \dfrac{1}{\left[\dfrac{X_3}{2X_2}\right]\left[\dfrac{R_0}{k_3^2(1 + \mu_2)}\right] + \dfrac{X_3(1 + k_2 - 2k_3^2)}{2k_3^2} + \left[\dfrac{1 - \mu_1}{1 + \mu_2}\right]\left[\dfrac{X_3}{4k_3^2 X_2}\right]\left[\dfrac{2SZ_N}{Z_N + 2S}\right]}} \right| \quad \text{Eq. (12)}$$

From Equation (12) the equivalent circuit of Fig. 9(a) can be developed. Starting at the lower right-hand side of Equation (12) it will be observed that $2SZ_N/(Z_N + 2S)$ represents Z_N and $2S$ in parallel. This parallel combination is multiplied by $(1 - \mu_1)X_3/(1 + \mu_2)(4k_3^2 X_2)$. In series with the parallel combination is the leakage inductance of transformer T equal to $X_3(1 + k_2 - 2k_3^2)/2k_3^2$; and the term containing R_0, which stands for $R_p + R_2(1 + \mu_2)$. In parallel with all of this is X_3. The resistance R_6 is simply a series resistance as evident from Fig. 8.

It can be seen from Fig. 9(a) that $X_3{:}4X_2k_3^2$ is the impedance ratio of transformer T, and that $-(\mu_1 - 1)/(\mu_2 + 1)$ is the transformation ratio of an ideal converter. The resulting circuit can be represented by Fig. 9(b).

Next consider the impedance Z_{34}, which is seen looking into terminals 3 and 4 of Fig. 8.

$$Z_{34} = \frac{E_4}{I_4} - Z_N \quad \text{Eq. (13)}$$

$$I_4 = \frac{\begin{vmatrix} P & Q & -M_3(1 + \mu_2) \\ Q & P & -M_3(1 + \mu_2) \\ -M_3 & -M_3 & Z_L + R_6 + X_3 \end{vmatrix} E_4}{\Delta} \quad \text{Eq. (14)}$$

$$\frac{E_4}{I_4} = \frac{(Z_N + 2S)[(Z_L + R_6 + X_3)(P + Q) - 2M_3^2(1 + \mu_2)] - 2(1 - \mu_1)S^2(Z_L + R_6 + X_3)}{(P + Q)(Z_L + R_6 + X_3) - 2M_3^2(1 + \mu_2)} \quad \text{Eq. (15)}$$

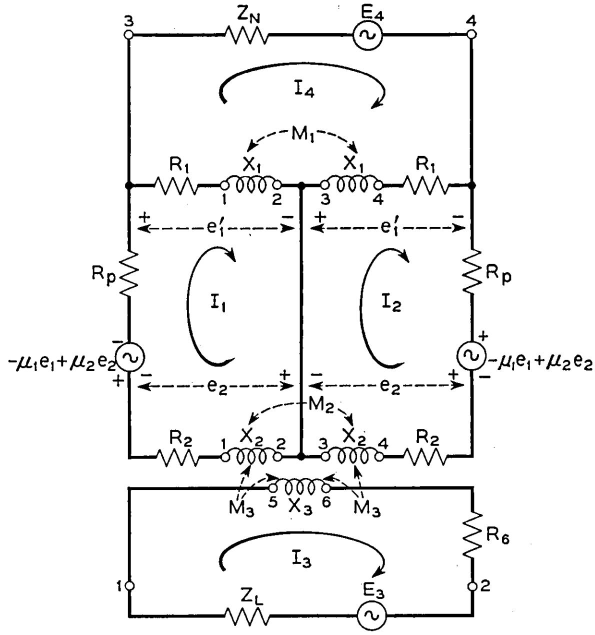


Fig. 8—Schematic of E1 repeater.

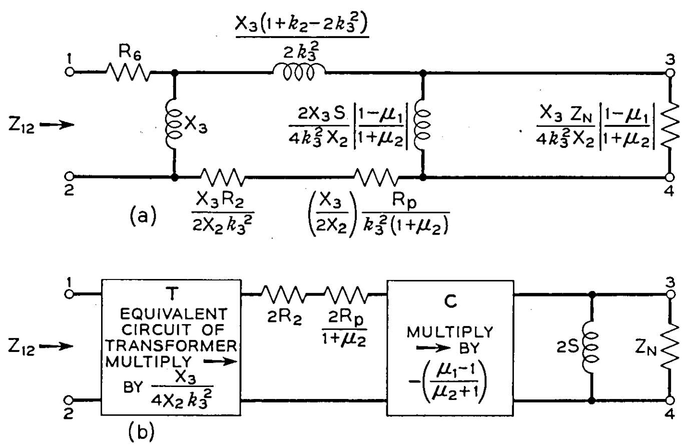


Fig. 9—Equivalent circuit of equation (12).

$$Z_{34} = 2S - \frac{2(1 - \mu_1)S^2(Z_L + R_6 + X_3)}{(P + Q)(Z_L + R_6 + X_3) - 2k_3^2X_2X_3(1 + \mu_2)} \qquad \text{Eq. (16)}$$

Substituting for $P + Q$ and rearranging

$$Z_{34} = \left| \frac{1}{\dfrac{1}{2S} + \dfrac{1}{\dfrac{2R_p}{(1-\mu_1)} + \dfrac{2X_2(1+\mu_2)}{(1-\mu_1)}[1 + k_2 - 2k_3^2] + \dfrac{4k_3^2X_2}{X_3}\dfrac{(1+\mu_2)}{(1-\mu_1)}\dfrac{X_3(Z_L+R_6)}{(Z_L+R_6+X_3)}}} \right| \qquad \text{Eq. (17)}$$

From Equation (17) the circuit of Fig. 10(a) can be developed in a manner similar to the way Fig. 9(a) was developed from Equation (12). Furthermore, Fig. 10(a) can be rearranged in the form shown in Fig. 10(b).

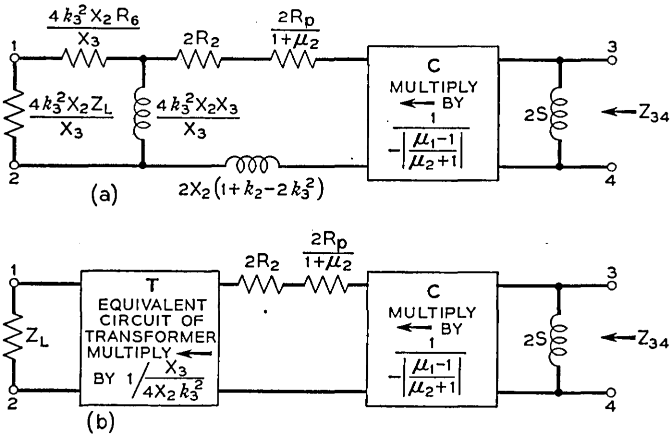


Fig. 10—Equivalent circuit of equation (17).

A comparison of Fig. 9(b) with Fig. 10(b) shows these two circuits to be essentially the same except for the point of view. If it is remembered that Z_N in Figs. 8, 9 and 10 has been chosen to include the elements in the RC network which couples the plate of one vacuum tube to the grid of the other, then from consideration of Figs. 9(b) and 10(b) it is obvious that Fig. 6(b) is the equivalent circuit of Fig. 6(a).

If Equation (6) or Equation (7) were studied it would be found that for all conditions of circuit stability the current I_3 would flow in the same general direction in which it would be expected to flow in any circuit mesh where all impedances were positive. But Z_{12} can equal a negative impedance. Hence, Z_{12} must be a reversed voltage type of negative impedance equal to $-V/I$. For an impedance to be negative either the voltage V at some frequency must be the reverse in polarity of that measured across a like positive impedance or the current must flow in the reverse direction. If the current does not reverse direction the voltage must, to produce a negative impedance.

If Equation (14) for the current I_4 were studied it would be found that all conditions which meet the requirement for circuit stability when the impedance Z_{34} is negative would produce a flow of current through this mesh with a phase impossible to realize were Z_{34} positive. Hence, if impedance Z_{34} is to be negative the current must be reversed and the impedance must be of the reversed current type equal to $V/-I$, where here V is the voltage across Z_{34} and I is the current flowing through it.

Design Considerations

If a resistance were connected to the network terminals 3 and 4 of Fig. 6(a) the impedance seen at the line terminals 1 and 2 would follow a locus

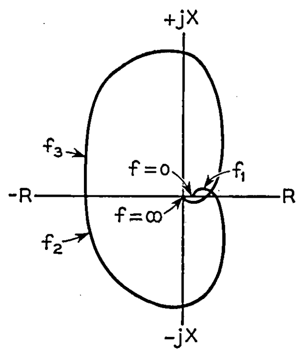


Fig. 11—Impedance locus of E1 repeater.

similar to that shown in Fig. 11 for the frequency range from zero to infinity. At zero frequency this impedance would be a small positive resistance equal to the d-c resistance of the primary windings of transformer T. At some low frequency f_1 (Fig. 11) the locus would show a positive impedance. In the E1 repeater it is this portion of the impedance characteristic which is used for the passage of low frequencies such as ringing, dialing and the like. Between the frequencies of f_2 and f_3 (Fig. 11) is seen an impedance which approximates a negative resistance. At high frequencies the locus would approach the origin again. In Fig. 11 this approach is shown through the first and then the fourth quadrant. At frequencies above the speech band passed by the telephone line a negative impedance is not wanted for the E1 repeater because it is of little value for voice transmission and may be detrimental in adding to the difficulty of obtaining stable operation.

In the design of the E1 repeater it is desirable for the network to have control of the impedance presented at the line terminals over the voice frequency range (f_2 and f_3 of Fig. 11). To accomplish this all impedances shunting the converter circuit are made as large as possible and all impedances in series are made as small as possible at these frequencies. Furthermore, μ_1 is made as close to μ_2 as practical, and large enough so that the factor $(\mu_1 - 1)/(\mu_2 + 1)$ is made as close to unity as possible. If this factor approximates unity and if the series term $R_p/(1 + \mu_2)$ is relatively small with respect to impedances between which the converter is operating, it can be seen from the equations that battery supply variations and tube changes should have little effect upon the negative impedance presented by the converter at these frequencies.

The retard coil and transformer inductances should be considered at low frequencies as a possible source of instability. At low frequencies the inductance of the retard coil shunting terminals 3 and 4, and the inductance of the transformer shunting terminals 1 and 2 materially affect the impedances facing the ideal converter.

Another important consideration in the design is the ratio of the line transformer. The ratio must be such that the tubes will operate efficiently with the normal network, line load, and plate supply voltages. This part of the design follows conventional methods.

The distributed capacitances of the line transformer windings should be taken into account. In general, these are large compared to the tube capacitances between cathodes and between ground and each cathode. These interelectrode capacitances can be neglected at voice frequencies, but at the higher frequencies all of them should be considered, and conventional methods of suppression of parasitic oscillations should be applied. Also, the distributed capacitance of the retard coil must be considered from the standpoint of stability at the higher frequencies.

Conclusion

A vacuum tube arrangement for producing negative impedance can be represented by an equivalent four-terminal network consisting of an ideal converter having a ratio of transformation of $-k{:}1$ and two positive impedance networks located one on each side of the ideal converter. In these two networks some of the elements can be cancelled in effect by balancing those in one network against corresponding elements in the other. Elements which cannot be balanced can be made either relatively large or relatively small compared to the impedances between which the converter is designed to operate. Across one pair of the four terminals of a practical converter can be seen a reverse voltage type of negative impedance; across the other pair can be seen a reversed current type. This device can be made to approach the ideal only over a finite frequency range. At some frequencies k will not be a numeric. Likewise, at some frequencies the two internal networks will produce appreciable effect. Stability can be ascertained readily from a working knowledge of the components of the equivalent circuit. Designs are practical wherein variations in battery voltages and vacuum tube constants are second-order effects.

Acknowledgements

The writer acknowledges the help of Messrs. K. G. Van Wynen, K. S. Johnson and F. B. Llewellyn with the concept of negative impedance and the contributions of Messrs. J. A. Weller, H. Kahl and W. J. Kopp relating to the practical design of negative impedance converters. He is also indebted to Miss A. B. Strimaitis for the many computations which were required in this development and to Messrs. R. Black, J. A. Lee and J. M. Manley for their comments on this paper.

In recent years there has been a growing need for an inexpensive two-wire repeater having a cost lower than that of the 22 type for use in the exchange plant. Mr. G. C. Reier encouraged the study of negative impedance devices as a possible way of meeting this need. Had it not been for his encouragement the E1 repeater development might not have been undertaken.

The Ring Armature Telephone Receiver

By E. E. MOTT and R. C. MINER

(*Manuscript Received Aug. 15, 1950*)

A new type of telephone receiver is described, in which the permanent magnet, the pole piece and the armature, which drives a light weight dome, are all ring-shaped parts. This structure exhibits a substantially higher grade of performance than present receivers of the bipolar type, with regard to efficiency, frequency range, leakage noise level, and response when held off the ear. In addition to showing the characteristics of this new receiver, an analysis of the various losses is given, and ideal performance limits are established. The advantage of providing an auxiliary path for the air gap flux is indicated, and other applications of the device as a transducer are described.

INTRODUCTION

THE ring armature receiver is a new type of telephone receiver developed for use in the subscriber's telephone set. It differs from other types in that the diaphragm consists of a thin, lightweight, dome-shaped central portion made of low density, non-magnetic material whose function is to radiate sound energy, surrounded by a narrow ring-shaped armature to which it is attached. The ring armature is not clamped at the outer periphery, but is held in place solely by magnetic attraction. It is driven at its inner periphery by the magnetic force. A ring-shaped pole and magnet structure serves as the motor element to drive the diaphragm. The new receiver is shown in sectional view in Fig. 1.

The advantage of the composite diaphragm construction in the new receiver is that the central portion moves almost wholly like a piston and is therefore nearly 100% effective and that its contribution to the total moving mass is small, being of the order of $\frac{1}{5}$. For these reasons it has been found possible to reduce the mass per unit area to approximately $\frac{1}{3}$ of that of the diaphragm of the bipolar receiver. Because of the large effective diaphragm area and the low mass, the acoustic impedance of the new receiver is low, being about $\frac{1}{7}$ of that of the earlier receiver. Although the motor efficiency is approximately equal to that of the bipolar receiver, the improved diaphragm construction yields a receiver of higher available power response,[1] wider frequency range, improved characteristic when the receiver is held off the ear, and having greater discrimination against room noise.

The ring armature construction is also applicable to devices other than earphones, such as microphones and loudspeakers.

[1]A.S.A. Standard Z24.9-1949 "Coupler Calibration of Earphones."

EARLY STEPS IN THE DEVELOPMENT

It has long been evident that the effective mass of the magnetic disc type diaphragm used in the bipolar receiver is high, and is therefore a serious limitation to obtaining an extended frequency range without sacrificing efficiency. Attempts were made to reduce this mass by using lightweight cone type radiators driven by relatively small magnetic discs at the center as in Fig. 2(a). The difficulties, however, of controlling the vibrational stability

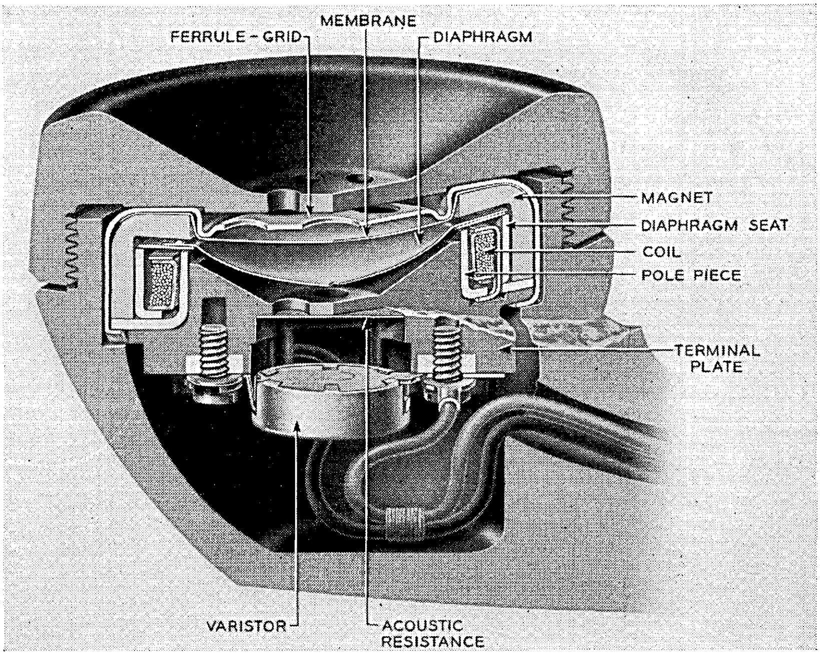


Fig. 1—Sectional view of the ring armature receiver in a handset.

of such structures made them impractical. Moreover, the mass of the armature was 100% additive to the moving system.

By reversing the positions of the armature and the light-weight portion of the diaphragm, putting the latter in the middle and using a ring of magnetic material around the outside edge as shown in Fig. 2(b), much better results were obtained.[2,3] The armature, having one edge resting on a seating surface, added only 30% of its mass to the moving system. Also, since the large central portion of the diaphragm carries no flux, it could now be replaced by a lightweight non-magnetic material instead of the relatively

[2]U. S. Patent No. 2,170,571, E. E. Mott, Filed August 12, 1936.
[3]U. S. Patent No. 2,171,733, A. L. Thuras, Filed October 6, 1937.

heavy magnetic material needed for armatures. These two factors permitted a very substantial reduction in the effective mass to be made in the moving system. In addition it has been observed that the peripherally driven diaphragm moves as a piston over a wide frequency range, while a centrally driven cone type diaphragm, as in Fig. 2(a), is more likely to have parasitic modes of vibration in the frequency range of interest. However, the receivers of the type shown in Fig. 2(b) needed small air gaps in order to get the force factor necessary to attain a high efficiency. Moreover, the thin

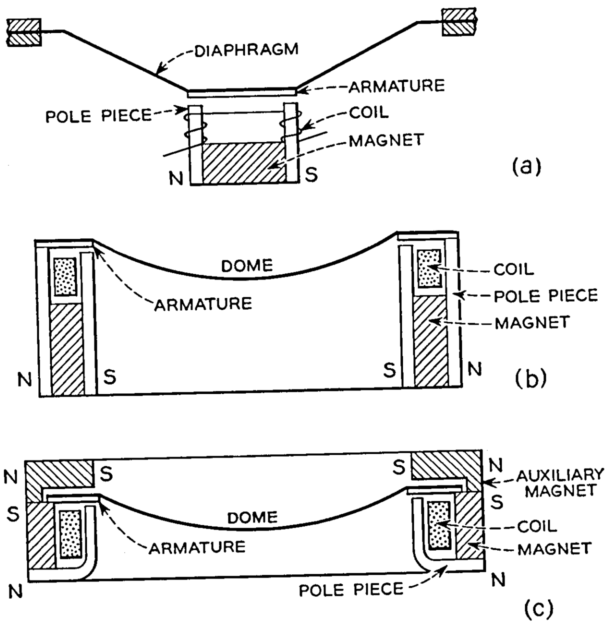


Fig. 2—(a) Early composite diaphragm receiver.
(b) Simple ring armature receiver.
(c) Ring armature receiver with auxiliary magnet.

magnet mounted between the inner and outer pole pieces presented manufacturing problems because of curved surfaces that had to be ground to close tolerances. The magnet in this case had to be of low reluctance and large section area, and this resulted in a rather tall structure.

By adding an auxiliary ring magnet overlying the front of the diaphragm, as in Fig. 2(c), radially magnetized in aiding relation to the lower magnet, a large increase in force factor was attained with larger air gaps than in Fig. 2(b).[4, 5] In addition an upright main magnet, as shown, could then be

[4]U. S. Patent No. 2,249,160, E. E. Mott, Filed May 19, 1939.
[5]U. S. Patent No. 2,249,158, L. A. Morrison, Filed July 15, 1941.

used to advantage, even though it had a relatively higher reluctance. With this design, magnetic fields were produced in the two air gaps above and below the diaphragm. In addition, the auxiliary ring magnet acted to shunt a portion of the d-c. flux around the armature, which resulted in reduced saturation in the middle portion of the armature, which permitted increased flux density in the air gap below the diaphragm. This partially separated the paths of the a-c. and d-c. flux in the magnetic structure, and adjusted the magnetic forces exerted on the diaphragm, so that lower stiffness armatures

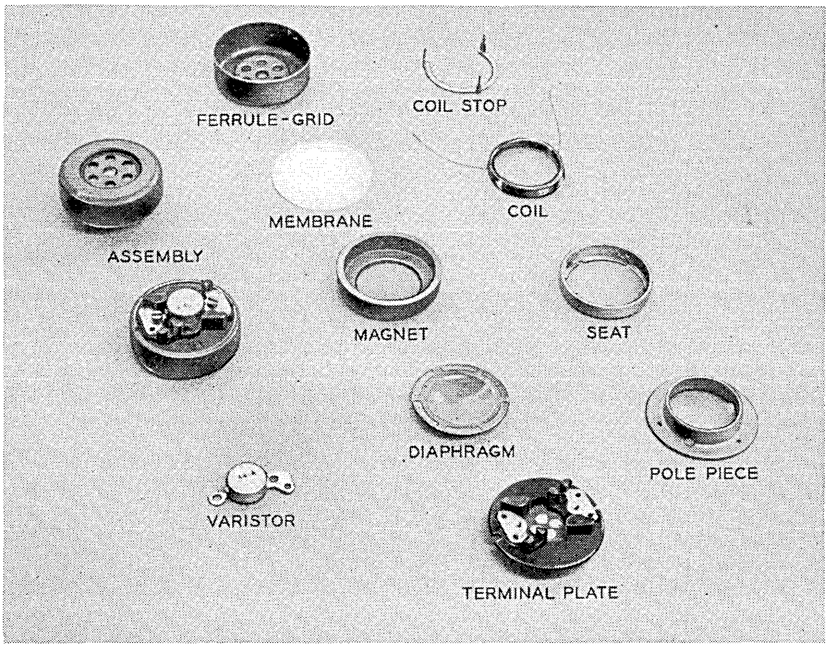


Fig. 3—Photograph of the ring armature receiver and its parts.

could be used than in the structure of Fig. 2(b). Application of the auxiliary ring magnet in front of the diaphragm contributed very substantially toward the development of a suitable motor element. The coil and magnet relationship shown in Fig. 2(c) also resulted in a lower axial height and a more compact structure.

Description of the Production Design

The main and auxiliary magnets of the early designs were Alnico castings, and were expensive to produce. In the production design, they have been combined into a single L-sectioned remalloy ring magnet, and the diaphragm seat has been transferred from the main magnet to a non-magnetic nickel-chromium alloy ring.

The details of this construction are shown in Figs. 1 and 3. The diaphragm consists of a flat ring-shaped vanadium permendur[6] armature, to the inner periphery of which is cemented a lightweight, plastic-impregnated and molded cloth dome. The dome is placed in a concave downward position with respect to the receiver cap, which greatly enhances its strength with respect to suddenly applied air pressures. At the outer periphery, the armature rests on the non-magnetic seat, and is driven at the inner margin by the magnetic field produced at the tip of the ring-shaped pole piece. The 45% permalloy[6] pole piece, at its end opposite the air-gap, has an integral flange extending outward, to which the armature seat is welded. The magnet also rests on the pole piece flange, outside of the armature seat. This magnet is made of a material similar to remalloy,[6] but of a modified composition in which the molybdenum content is increased, which results in a higher coercive force but lower residual induction. It is hot formed from sheet material into the shape of a cup and then punched to provide an opening in the base of the cup, the diameter of the opening being slightly smaller than the inner diameter of the cylindrical portion of the pole piece. The inwardly extending flange of the magnet forms the auxiliary portion.

Acoustic controls of the response-frequency characteristic of the receiver are provided in the same manner as in former telephone receivers of the controlled-diaphragm type,[7] except that lower values of acoustic impedance are used which are easier to control. Coupling chambers on each side of the diaphragm connect to constricted passageways having acoustic mass and resistance. The coupling chamber under the diaphragm exhausts into the handset handle cavity through four holes, molded in the phenol plastic terminal plate, which are covered with an acoustic resistance fabric cemented to the terminal plate. This acoustic mesh serves to extend the frequency range of the receiver because of its negative reactance characteristics, and the acoustic resistance damps out the diaphragm resonance. The coupling chamber above the diaphragm exhausts into the listener's ear cavity through the acoustic mass and resistance of the holes in the receiver cap. Proper selection of the acoustic impedances of the elements of this mesh serves still further to extend the frequency range of the receiver. The relationships of all the acoustic and mechanical elements is such as to produce the desired response frequency characteristic (See section entitled "Network Representation").

There are several parts of the ring armature telephone receiver whose

[6] "Survey of Magnetic Materials and Applications in the Telephone System," V. E. Legg, *The Bell System Technical Journal*, Vol. XVIII, July 1939.

[7] Instruments for the New Telephone Sets, W. C. Jones, *The Bell System Technical Journal*, Vol. XVII, July 1938.

functions are almost completely mechanical, as compared with the magnetic, electrical, or acoustical functions of other parts. The armature seat, which already has been mentioned, is one of these. An interesting design feature of this part is the necessity for high electrical resistivity since most of the a-c. flux links the seat and it is therefore subject to eddy current losses. A nickel-chromium alloy has been found suitable for this part. Another part having a purely mechanical purpose is the coil stop. This part consists of a flat strip of metal punched in a curved shape so as to fit on top of the coil, and having three prongs bent at right angles to the strip. The tips of the prongs pass through slits in the pole-piece flange and are bent over in assembly to hold the coil in place. A membrane is mounted between the protective grid and the magnet flange to keep dust and other foreign substances out of the instrument. In this receiver the protective grid and the clamping ferrule, which is crimped over in the final assembly, are combined into one part.

Low manufacturing costs are realized by the use of multiple-purpose parts. Some examples have been noted already. The single magnet serving as both main magnet and auxiliary magnet is an example. The combined ferrule-grid, which eliminates the fabrication, finishing, and handling of one part as compared with previous designs, is another. The terminal plate also falls into this class of parts. It not only serves as an electrical termination for the receiver, but also is molded in such a way as to provide the correct coupling air volume in back of the diaphragm; it contains the acoustic passageways leading out of the back of the instrument and provides a mounting surface for the acoustic resistance fabric cemented over these passageways; it mounts and protects a click-reduction varistor which is made a part of the receiver; it is molded with projections which prevent the spade tip terminals of the handset cord from shorting against the varistor case or turning in such a manner as to cause the cord conductors to be pinched between the receiver and its handset seating surface; it has other projections which key into the coil lead holes of the pole-piece and provide insulation between the wires and the pole-piece and at the same time orient and prevent rotation of the terminal plate with respect to the pole-piece; and it is provided with two slots into which the crimped edge of the ferrule is staked to prevent rotation of the ferrule.

Perhaps the most interesting example of a multiple-purpose part is the diaphragm dome. Its primary purpose is, of course, to radiate sound energy in its capacity as a lightweight, rigid closure for the central opening of the armature. In addition, it has six small projections molded to its top surface just outside of the dome portion, which will touch the inner edge of the magnet flange if the diaphragm is lifted upward off of its seat by mechanical shock. Thus the projections prevent the armature from coming close enough to the auxiliary magnet to be held there by magnetic attraction and insure that the armature will always return to its seat in case it is dislodged by mechanical shock. Another function which is built into the dome serves to prevent the outside edge of the armature from coming into contact with the inside surface of the cylindrical main portion of the magnet. Contact of this nature has been found to produce irregularities in the response-frequency characteristic of the receiver and a variability of the output level of a few decibels. These undesirable variations are controlled by six spokes of the dome material which extend outward beyond the edge of the armature a few thousandths of an inch and thus tend to prevent contact between the armature and magnet. Finally, the dome contains a small hole which introduces a low frequency cut-off in the response-frequency characteristic of the re-

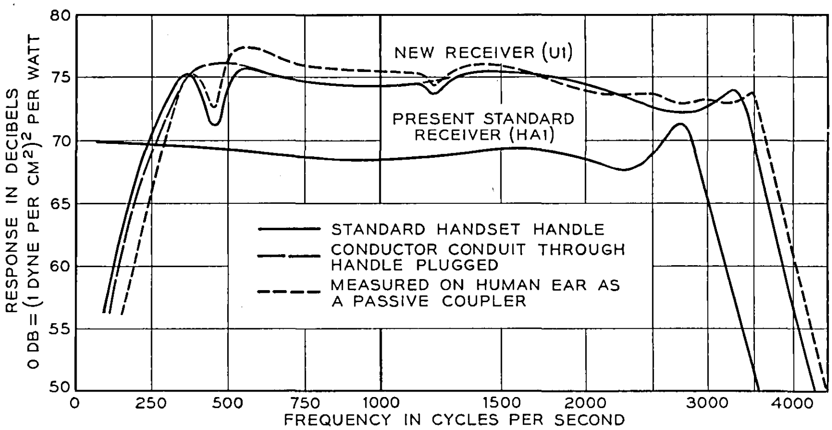


Fig. 4—Available power response-frequency characteristics measured with a source impedance of 128 ohms on a 6 cc. coupler, except as noted.

ceiver, which has been found to be desirable to reduce interference picked up in telephone circuits from electrical power circuits.

Performance Characteristics

A partial evaluation of the performance improvements of the ring armature receiver as compared with its predecessor is illustrated in Fig. 4, which shows available power response-frequency characteristics[8] for the two receivers. The two solid curves in this figure show the relative sound pressure output of the new U1 receiver and the present standard HA1 receiver over

[8]Response-frequency characteristics in this article are shown on a new frequency scale which gives a well balanced visual emphasis to the various frequency bands. The scale is linear from 0 to 1000 cycles per second, and logarithmic from 1000 to 10,000 cycles per second, the two sections having a dimensional ratio of 4 to 9. See "A New Frequency Scale for Acoustic Measurements", W. Koenig, *Bell Laboratories Record*, August 1949.

their frequency ranges. These solid curves were measured with the receiver on a standard closed coupler[9] of six cubic centimeters volume, using a source impedance of 128 ohms, and the ordinate scale is given in terms of the square of the pressure generated in the coupler per unit of electrical power available to a pure resistance of 128 ohms substituted in the electrical circuit in place of the receiver. The new receiver shows 5 decibels improvement in output level and about 500 cycles per second extension in the frequency range. This represents a very substantial increase in transducer efficiency, and the increase in range results in a quite noticeable improvement in the quality of speech sounds. The low frequency cut-off obtained by a hole in the diaphragm of the U1 receiver, mentioned in the preceding section, appears in the response-frequency characteristic below 350 cycles per second. The irregularities in the characteristic of the U1 receiver at 450 and 1200 cycles per second are not inherent in the receiver, but are acoustical effects of the passageway molded in the handset handle, which serves as a conduit for the wires connected to the receiver unit. This is indicated by the dashed line curve, which shows the response-frequency characteristic of the receiver when the passageway is plugged at the receiver bowl of the handset. No adverse effect of these irregularities has been discerned.

For comparison with the closed coupler characteristic, the dotted curve in Fig. 4 shows the pressure generated by the U1 receiver at the entrance to the human ear. This curve is an average of 90 observations on 30 subjects measured by a small diameter search tube inserted into the outer ear cavity through the receiver cap and connected to a microphone external to the handset, so that the ear is used as a passive coupler. Figure 5 shows the manner of using the apparatus, which includes a 640AA condenser transmitter mounted on the handset and coupled to the search tube through a very small chamber. It will be noted that the curve of Fig. 4 taken on a human ear shows increased low frequency cutoff because of leakage between the receiver cap and the ear, but that otherwise the 6 cc. closed coupler response is a good representation of the data taken on the ear. Considerable deviations from the average curve were observed from person to person, as illustrated in Fig. 6 which shows the maximum and minimum values of all measurements at various frequencies and the standard deviation of the measurements at three frequencies.

An interesting comparison of the performance of the U1 and HA1 receivers is made by holding the receivers slightly away from the ear. It is observed that the degradation in response caused by this condition, which represents a very large amount of acoustical leakage between the ear and the receiver cap, is much greater in the HA1 receiver. The effect is illus-

[9]Type 1 coupler per A.S.A. Standard Z24.9-1949.

trated by Fig. 7, which shows available power response-frequency characteristics for the two receivers when they are raised one-eighth of an inch from the normal sealed position on a standard coupler. The U1 receiver shows better response than the HA1 receiver at both high and low frequencies. The low frequency end is cut off less in the U1 receiver, because it has 2.5

Fig. 5—Method of measuring receiver response on a human ear, using a search tube microphone.

times larger effective area and therefore is a better radiator of sound at low frequencies. The high frequency end is better because of the inherent extension of frequency range in the U1 receiver.

Another interesting characteristic of the new receiver is its performance under conditions of high ambient noise levels. Noise leakage between the receiver cap and the external ear generates an acoustic noise pressure in the ear cavity which may mask the sound signal from the receiver. This leakage

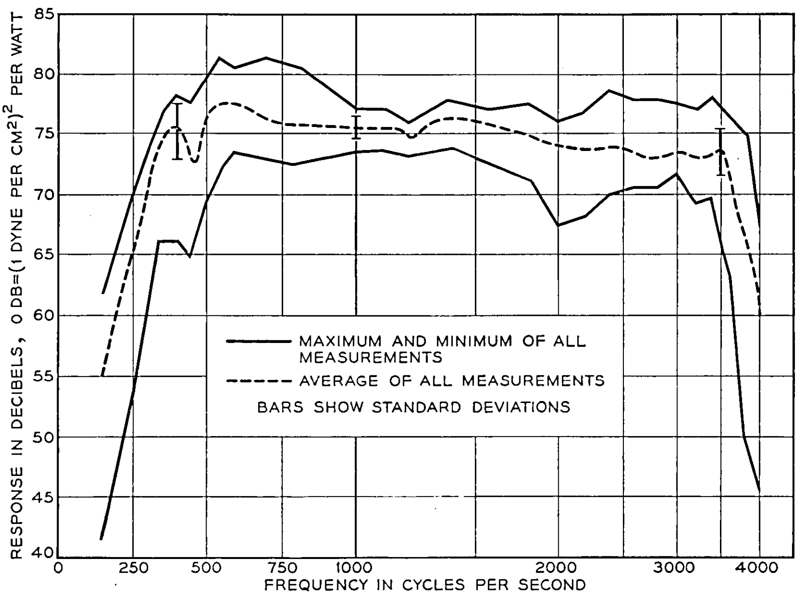


Fig. 6—Available power response-frequency characteristics of U1 receiver, measured with a source impedance of 128 ohms, on human ears as passive couplers.

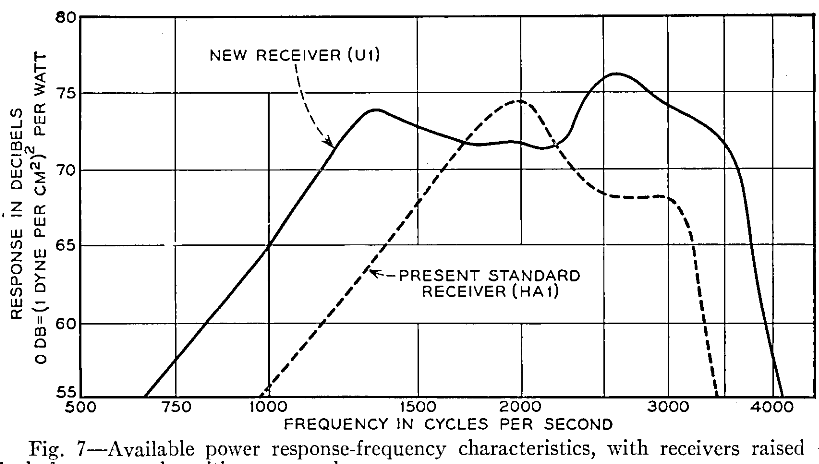


Fig. 7—Available power response-frequency characteristics, with receivers raised $\frac{1}{8}$ inch from normal position on coupler.

noise is predominantly low frequency noise not only because typical room noise characteristically decreases with increase in frequency,[10] but also because the leakage path has an acoustic impedance which rises with fre-

[10] "Room Noise Spectra at Subscribers' Telephone Locations," D. F. Hoth, *Journal of Acoustical Society of America*, Vol. 12, April 1941.

quency. The U1 receiver has considerably lower acoustic impedance than the HA1 receiver and, since the acoustic impedance of the receiver shunts the ear coupling chamber impedance, the same degree of noise leakage under the U1 receiver generates a lower noise pressure in the ear cavity with a resultant decrease in the masking effect on a given signal. Figure 8 shows data on this characteristic. Tests were made by measuring the pressure in a 6 cc. coupling chamber closed by the receiver except for a leakage path having acoustic resistance and mass values approximating those of the worst leakage condition shown by the data presented in Fig. 6. The measurements were made in a highly absorbent room, with a loud-speaker as the sound source. The curves show that below 1500 cycles per second the noise leakage sound pressure generated in the coupler when the U1 receiver is used is less than that of the HA1 receiver by as much as 5 db over a considerable portion

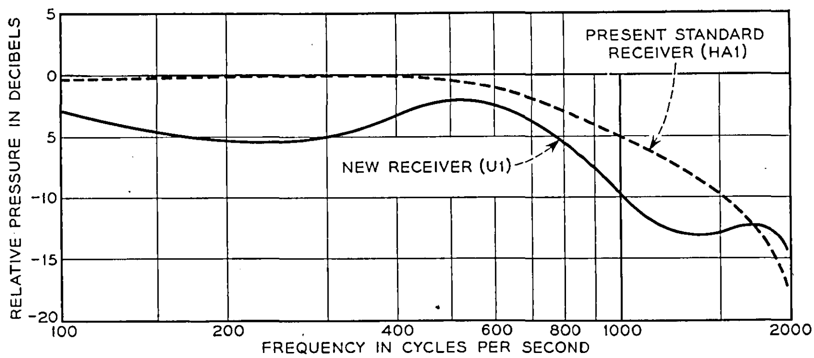


Fig. 8—Relative noise leakage sound pressure in a 6 cc. coupler having a simulated ear leak. Reference level is the pressure at the coupler with the receiver removed.

of the frequency range. The effect of this difference has been observed in listening tests using real voices* and in subjective tests† while measuring the shift in the threshold of intelligibility between the two receivers under quiet and noisy conditions. Since under certain conditions a reduction in the noise may be equivalent to a corresponding gain in signal strength, this feature of the ring armature receiver represents a distinct improvement.

A characteristic of considerable importance in the development and design of a telephone receiver is the manner in which the receiver output level varies with direct current superimposed on the alternating current flowing in the receiver coils. The direct current may be applied in such a way that it develops flux in the magnetic circuit which either aids or opposes the polar-

*Unpublished work by W. D. Goodale, Bell Telephone Laboratories.
†Unpublished work by R. H. Nichols, Bell Telephone Laboratories.

izing flux of the permanent magnet. Superimposed direct-current characteristics are shown for both the U1 and the HA1 receivers in Fig. 9. The vertical line in the plot labeled zero represents the normal operating condition of the receiver with no direct current in the coils. Increasing values of opposing current are plotted to the left and increasing values of aiding current are plotted to the right of this axis. In order that such curves may be compared fairly, they have been plotted on the basis of equal impedances for the two instruments, and each receiver has been referred to an impedance of 100

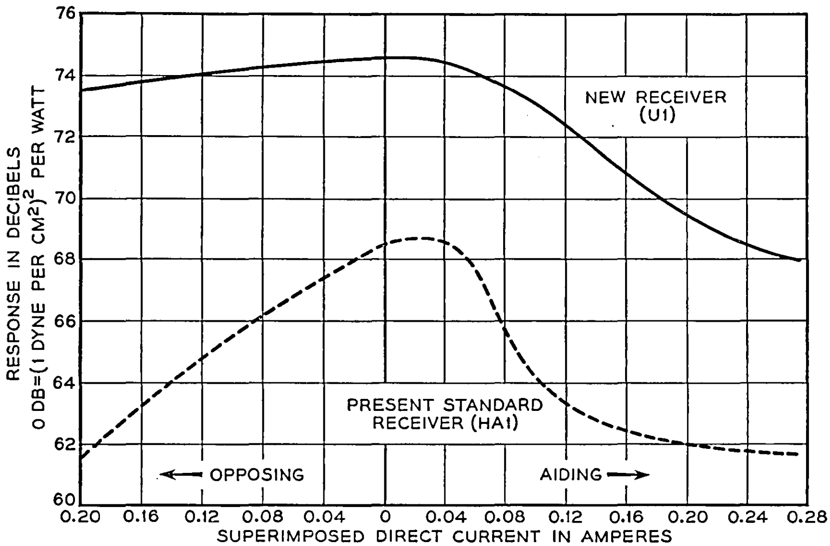


Fig. 9—Available power response versus superimposed direct current characteristics measured from a source impedance of 128 ohms on a 6 cc. coupler at 1000 cps. Each receiver has been referred to 100 ohms impedance by multiplying its current values by $\sqrt{\frac{Z_{1000}}{100}}$.

ohms. Thus, if both receivers were of 100 ohms impedance at a frequency of 1000 cycles per second the curves would be compared directly. If an instrument has an impedance differing from 100 ohms, its direct current values along the scale are multiplied by a suitable scale factor as follows:

$$\text{Current scale factor} = \sqrt{\frac{Z_{1000}}{100}}$$

where Z_{1000} is the magnitude of the receiver impedance at 1000 cycles per second.

The superimposed direct-current characteristics are an indication of the stability of the receivers. Slight changes in the air-gap of a receiver may occur during its life due to external mechanical stresses, extreme temperature variations, or magnetic influences. The application of direct current to the receiver winding produces such changes in the air-gap artificially in a controllable and reproducible manner, and shows that within a certain range there is no serious effect on the response level to be expected from slight alterations in the air-gap. The curve also shows that there is an optimum magnetization for the instrument at which the response is a maximum, and the sharpness or bluntness of the response peak is a measure of the stability.

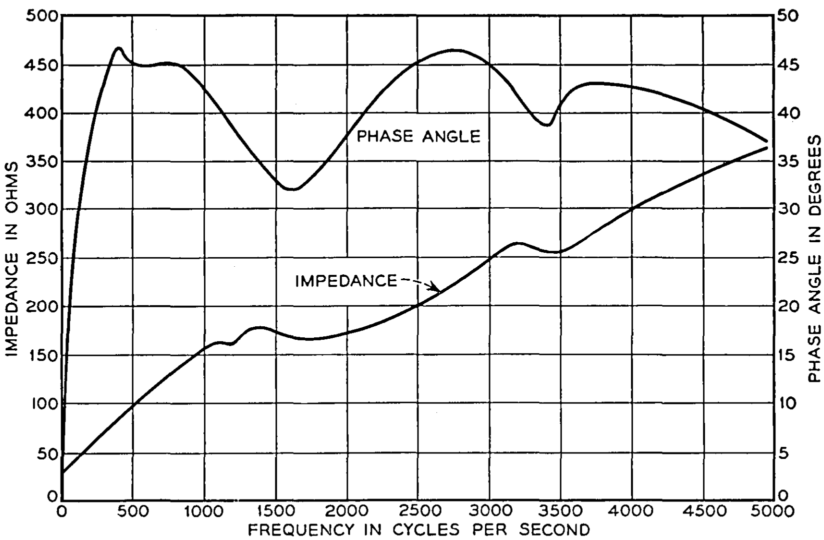


Fig. 10—Impedance of ring armature receiver, measured on a closed chamber of 6 cc volume.

The comparison between the two receivers shows that the U1 receiver is less sensitive to changes in superimposed direct current than the HA1 receiver.

The impedance-frequency characteristic of a typical U1 ring armature receiver, measured on a closed chamber of 6 cc. volume, is shown in Fig. 10. The magnitude of the impedance rises with frequency because of the inductance of the instrument. The departures of the impedance curve from a smooth upward sweep with frequency represent the contributions of the mechanical and acoustical elements of the receiver to the electrical impedance, that is, the motional impedance. The phase angle of the ring armature receiver averages 40°, which is somewhat lower than that of bipolar receivers such as the HA1 because of the larger part played by eddy currents in the impedance of the former.

Network Representation

The representation of electro-acoustical transducers as electrical networks has long been a useful tool.[11, 12] Extensive use of this analogy has been made in the development and design of the ring armature receiver. The saving of time and increase in accuracy and completeness of analysis possible with this technique is apparent when it is realized that a complete family of response-frequency characteristics, showing the effects of variation of one or more of the mechanical or acoustical constants of the instrument, can be obtained by electrical measurements of voltage on an electrical network for various settings of variable inductances, capacitances, or resistances which simulate the mechanical or acoustical constants of interest. The amount of work required to obtain the same information by building and testing mechanical and acoustical models is such that in many cases it would be impractical or impossible within a reasonable time interval.

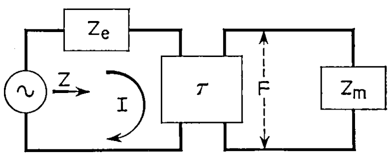


Fig. 11—Block diagram network representation for receivers.

A complete generalized representation of the ring armature receiver is shown in Fig. 11 in block diagram form. Z_e is the electrical impedance of the instrument with all motion of the armature blocked mechanically. Z_e includes a mesh which simulates the effects of eddy currents in the metallic structure of the instrument. τ is the force factor, defined as the force applied to the armature per unit of current flowing in the receiver winding. Z_m is the mechanical impedance of the mechanical and acoustical portion of the receiver at the point of application of the force, F. The relationships in this diagram are

$$\tau = \frac{F}{I} \quad \text{and}$$

$$Z = Z_e + \frac{\tau^2}{Z_m}.$$

The term $\frac{\tau^2}{Z_m}$ is the motional impedance of the instrument.

[11]"High Quality Recording and Reproducing of Music and Speech," J. P. Maxfield and H. C. Harrison, *Bell System Technical Journal,* July 1926.
[12]"Theory of Magneto-Mechanical Systems as applied to Telephone Receivers and Similar Structures," R. L. Wegel, *Am. Inst. of Elec. Eng.,* October 1921.

In studying the performance of an instrument the three elements in the above block diagram, Z_e, τ, and Z_m are considered separately, and a differ-

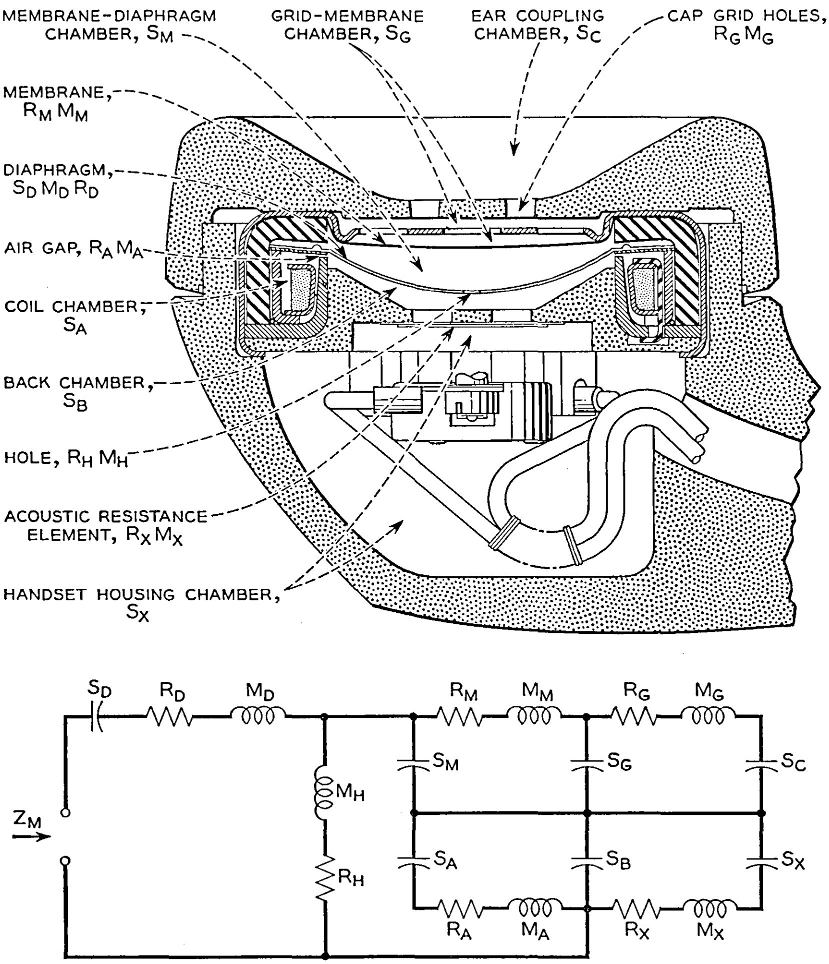


Fig. 12.—Z_m. Equivalent network of mechanical and acoustical elements of ring armature receiver.

ent type of analysis applied to each. These three elements and their analyses and uses will be discussed separately below:

(a) *Mechanical and Acoustical Elements*

The element of the block diagram represented by Z_m is the portion of the network representation of most use in the development and design of receivers. Figure 12 shows an expansion of Z_m into an equivalent electrical network which has been used extensively in the study of the ring armature receiver. The cross-sectional drawing of the receiver and associated handset handle and cap is labeled to indicate the various acoustical and mechanical elements which are represented by the electrical circuit in the lower portion of the figure. Thus S_D, M_D and R_D represent the stiffness, mass, and mechanical resistance of the diaphragm; M_H and R_H represent the mass and resistance of the hole in the diaphragm which provides the low-frequency cut-off in the receiver, etc. One item of interest in this circuit representation, which differs from that of previous receivers, is the division of the chamber in back of the diaphragm into two parts, S_B and S_A, connected by the air passageway R_AM_A of the magnetic air-gap between the armature and the pole-piece tip. Under certain conditions, particularly those representing the receiver with some of the acoustical controls removed, the acoustical constants of the air-gap have been found to be of sufficient magnitude to warrant this division of the total back chamber into two connected parts. An approximation is involved in this representation of the back chamber in that the force applied to the coil chamber, S_A, by the motion of the armature is ignored, but this approximation is justifiable through a consideration of the relative magnitudes of the effective areas and volumes involved, and the representation has been found to be in good agreement with measurements on the actual physical structures.

The constants of the equivalent circuit are determined by various physical measurements and computations. For example, the effective mass of the diaphragm, M_D, is estimated from the weights and the integrated vibratory kinetic energy of its various parts. The diaphragm stiffness, S_D, is then computed from its resonant frequency. The diaphragm resistance, R_D, is determined from a circle diagram analysis. The various chamber stiffnesses are computed from their air volumes, knowing the integrated effective area of the diaphragm. The acoustical resistances and masses are obtained from a combination of theoretical computations and special tests on the network, using circuit conditions in which these constants play the predominant role.

In setting up this equivalent circuit for analysis and study, mechanical resistances are replaced by electrical resistances; masses are replaced by inductances; and compliances, the reciprocal of the stiffnesses, are replaced by capacitances. Response-frequency characteristics may then be determined by applying a constant voltage to the input terminals and noting the voltage across S_C, which is proportional to the pressure generated in the ear coupling chamber for constant force applied to the armature. The effects of changes in any of the elements can be determined by simply changing the electrical value of the equivalent network element and repeating the measurement. In this manner, optimum values or combinations of values may be determined to provide the desired response-frequency characteristic, the effect of proposed design changes may be predicted, and other characteristics determined without building large numbers of physical models.

(b) *Damped Electrical Impedance*

The damped electrical impedance of the receiver, that is, the electrical impedance when the armature is blocked so that it cannot move, is represented by Z_e in the block diagram of Fig. 11. The damped impedance of the ring armature receiver plotted against frequency is shown in Fig. 13. The

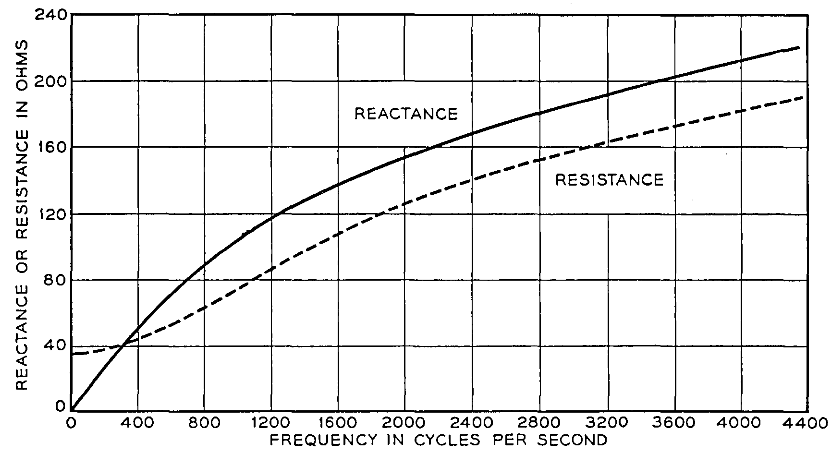


Fig. 13—Damped impedance of ring armature receiver without varistor.

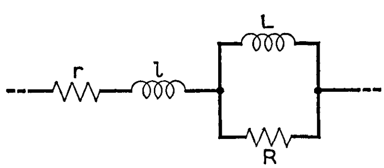


Fig. 14—Z_e. Network representation for the damped impedance.

rise in resistance and the departure from linearity of the reactance are due to the effects of eddy currents in the metallic parts of the instrument.

A circuit representation for Z_e is shown in Fig. 14. This circuit is derived on the assumption of a single eddy current path coupled to the receiver winding. The electrical resistance and inductance of the winding are represented by r and ℓ, and the eddy current circuit by R and L. Analysis of this circuit is useful in determining the extent to which eddy currents have a detrimental effect on the efficiency of the receiver. In general, the effect of eddy currents is greater in the ring armature receiver than in the bipolar types, largely because of the toroidal form of the motor element. Slotting of the ring-shaped parts has been found to be ineffective in reducing the eddy currents. However, with the low effective mass and large effective area of the diaphragm, the constants of the acoustical elements shown in Fig. 12 can be adjusted to compensate almost completely for the effects of eddy currents, even up to quite high frequencies.

(c) *Force Factor*

The third element in the block diagram of Fig. 11 is τ, the force factor, defined as the force on the armature per unit current flowing in the receiver

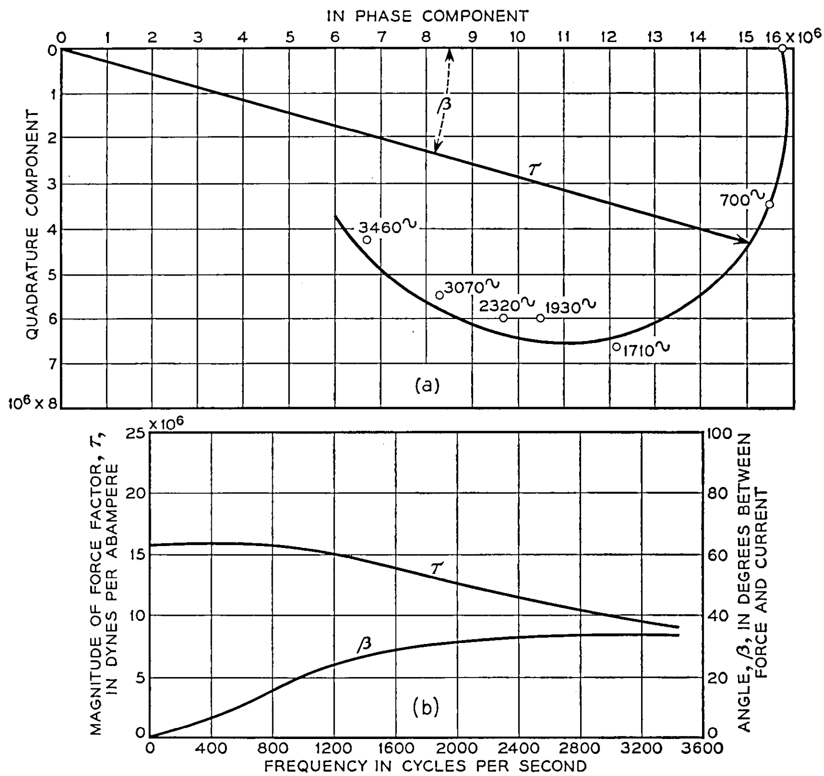


Fig. 15—Force factor plots of ring armature receiver. (a) Force factor circle. (b) Magnitude and angle of force factor.

winding. This is a complex quantity whose angle between force and current is designated by the symbol β. In magnetic receivers like the ring armature receiver, the force factor varies with frequency, both in magnitude and in phase. Fig. 15(a) shows a vector plot of this quantity, and indicates how the terminus of the vector follows an approximately circular path with change in frequency. Figure 15(b) is a chart showing the magnitude of τ and the angle β as functions of frequency.

The force factor diminishes with increasing frequency, and is largely dependent upon the alternating component of flux in the field of the air-gap. The presence of eddy currents produces a component of flux there also, which is usually in opposition to that produced by the current in the coil winding. The net amount of flux is thus diminished with increasing frequency as the effect of eddy currents becomes greater. It has been shown[12] that, for a single eddy current path, the locus of the vector plot such as that of Fig. 15(a) would be a semicircle with its center on the horizontal axis. The fact that the center of the circular locus shown here is somewhat below the horizontal axis indicates a departure from the simple theory. It seems likely, however, that if two or more eddy current paths exist having substantially different time constants, the departure shown here might be explained.

The values of force factor are obtained from a circle diagram analysis with the diaphragm resonated at different frequencies in the middle and upper frequency range. For very low frequencies, the mechanical impedance of the receiver is determined by the stiffnesses of the system, provided the hole in the diaphragm is closed, and a simple expression for the force factor at low frequencies can be derived in terms of the pressure generated by the receiver in a closed coupler for a given current in the receiver. This expression is

$$\tau_0 = \frac{(p_0)}{(I)} \cdot \frac{S_r + S_f}{S_f} \cdot A \text{ dynes per abampere.}$$

where τ_o is the low frequency force factor, p_o is the low frequency pressure generated in a closed coupler for current I in the winding, S_r is the total mechanical stiffness of the diaphragm and acoustic chambers back of the diaphragm referred to the effective diaphragm area A, and S_f is the mechanical stiffness of the closed coupling chamber in which p_o is measured.

Since the force per unit current will depend on the number of turns in the coil, force factor is not independent of receiver impedance. Thus, when comparing the force factors of receivers, it is necessary to refer the measured values of force factor to a common value of impedance by introducing a factor based on the ratio of the square roots of the impedances of the receivers being compared. Such comparisons between the ring armature receiver and its bipolar predecessor show approximately equal values of force factor at low frequencies, with the ring armature receiver force factor falling off more rapidly at higher frequencies.

Limits of Receiver Efficiency and Distribution of Losses at Low Frequencies

In the design of this receiver, an object has been to make the efficiency as high as practicable over the frequency range from 350 to 3500 cps. Since there are theoretical limits to the level of efficiency which cannot be surpassed even in the ideal case where there are no losses, it becomes of interest to determine the magnitude of these upper limits. It is also of interest to determine the amount and origin of the various types of losses that occur, and to what extent they can be minimized.

In the case of a telephone receiver, the character of the ear load is predominantly a reactance, corresponding to the stiffness reactance of an ear cavity of about 6 cc. volume. The power transfer to such a load is not ordinarily used to denote the efficiency, as may be done for a resistance terminated device such as a loudspeaker for example. Instead, the available power response is taken as a measure of relative efficiency of receivers, and it is defined as follows:

$$\text{Response} = 10 \log \eta = 10 \log \frac{|p|^2}{E^2/4R_0} \tag{1}$$

Where

$10 \log \eta$ = available power response in *db* referred to (1 dyne/cm^2)^2 per watt.

p = pressure developed in the ear cavity in dynes/cm^2.

E = voltage of source in volts.

R_o = source resistance in ohms; chosen in this discussion to be equal to the receiver impedance at the midband frequency f_o, 1000 cps.

In this expression, the numerator $|p|^2$ is proportional to the acoustic power output, while the denominator $E^2/4R_o$ is the available power input.

Low Frequency Loss Analysis

It will now be shown that the available power response approaches a theoretical limit which is about 17 *db* higher than the response level of this receiver at low frequencies, and that this expression may be arranged to give five loss factors, each of which has an important physical significance.

For this purpose we need to consider only the stiffness, force factor, and inductance of the receiver working into a closed chamber representing the ear load. It is well known that in this instance for frequencies below 500 cps, with the receiver working out of a source of constant voltage E, and resistance R_o, the equations of motion are:

$$\begin{aligned} (R_o + R + j\omega L)\, I + j\omega\tau\, .10^{-7}\, x &= E \\ -\, \tau I + (S_r + S_f)\, x &= 0 \end{aligned} \tag{2}$$

where L = low frequency inductance of receiver winding in henries

R = low frequency resistance of receiver winding in ohms

τ = transduction coefficient or force factor in dynes per ampere

S_r = stiffness of receiver diaphragm, including the rear chamber, negative field stiffness, etc. in dynes per cm.

S_f = stiffness load of coupler and front volume in dynes per cm.
I = current in amperes
x = displacement of diaphragm in cms.

The solution of these equations for the displacement is

$$x = \frac{\tau E}{(R_0 + R + j\omega L)(S_r + S_f) + j\omega\tau^2 \cdot 10^{-7}} \tag{3}$$

For an adiabatic change, the pressure p in the chambers in front of the diaphragm due to a small displacement x is

$$p = \frac{\gamma P_0 A}{V_g + V_c} x = \frac{S_f}{A} x, \tag{4}$$

where γ = ratio of specific heats of air = 1.405
P_o = atmospheric pressure = 1.013 x 10^6 dynes/cm^2
A = effective area of diaphragm in cm^2
V_g = receiver front volume beneath the grid holes, in cm^3
V_c = coupler volume = 6 cm^3

Combining equations (3) and (4) we have

$$p = \frac{\gamma P_0 A}{V_g + V_c} \cdot \frac{\tau E}{(R_0 + R)(S_r + S_f) + j\omega[\tau^2 10^{-7} + L(S_r + S_f)]}. \tag{5}$$

Substituting equation (5) in equation (1) and expressing η as a power ratio we obtain

$$\eta = \frac{\gamma P_0}{V_g + V_c} \cdot S_f \cdot \frac{4 R_0 \tau^2}{(R_0 + R)^2 (S_r + S_f)^2 + \omega^2[\tau^2 10^{-7} + L(S_r + S_f)]^2}. \tag{6}$$

Let $$K = \frac{\tau^2 10^{-7}}{L(S_r + S_f)} \quad \text{or} \quad \tau^2 = KL(S_r + S_f)10^7$$

Then

$$\eta = \frac{4\gamma P_0 .10^7}{V_c} \cdot \frac{V_c}{V_g + V_c} \cdot S_f \cdot \frac{R_0 K L (S_r + S_f)}{(R_0 + R)^2 (S_r + S_f)^2 + \omega^2[KL(S_r + S_f) + L(S_r + S_f)]^2} \tag{7}$$

and factoring

$$\eta = \frac{4\gamma P_0 .10^7}{\omega V_c} \cdot \frac{V_c}{V_g + V_c} \cdot \frac{S_f}{S_r + S_f} \cdot \frac{K}{1 + K} \cdot \frac{R_0 \omega L (1 + K)}{(R_0 + R)^2 + \omega^2 L^2 (1 + K)^2} \tag{8}$$

At low frequencies, the reactance of the receiver as seen from the electrical side is $X = \omega L(1+K)$. Hence the last factor of the above equation becomes $\dfrac{R_0X}{(R_0 + R)^2 + X^2}$. Taking the case of a pure reactance receiver first, we may place $R = O$, and if we further match R_0 to X at the midband frequency f_0, and then denote the value of X at f_0 as X_0, we have

$$\frac{R_0X}{(R_0 + R)^2 + X^2} = \frac{X_0X}{X_0^2 + X^2} = \frac{f/f_0}{1 + (f/f_0)^2}$$

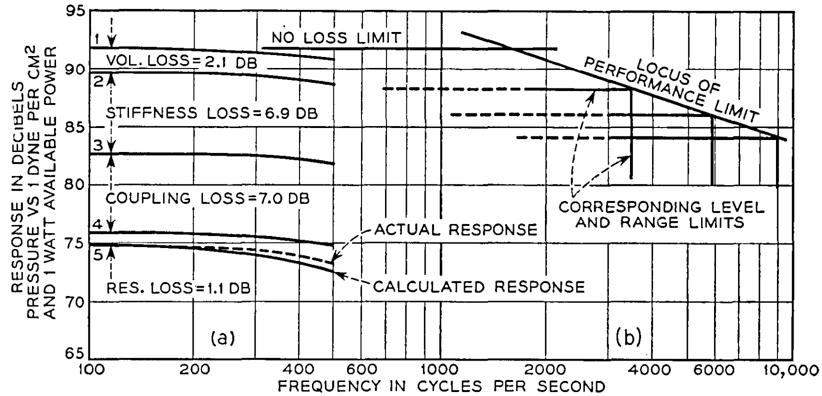


Fig. 16—(A) Low frequency loss distribution of ring armature receiver. (B) Theoretica limits of response level and frequency range, for receivers with uniform response down to zero frequency.

The latter step assumes that the low frequency reactance X is a linear function of frequency. For the pure reactance receiver, the response then becomes

$$10 \log \eta = 10 \log \frac{2\gamma P_0 .10^7}{V_c \pi f_0\left(1 + \dfrac{f^2}{f_0^2}\right)} \cdot \frac{V_c}{V_g + V_c} \cdot \frac{S_f}{S_r + S_f} \cdot \frac{K}{1 + K} \tag{9}$$

Thus it is seen that the expression for the efficiency may be factored into four terms, each of which has a significant physical interpretation.

A plot of these factors versus frequency is shown in Fig. 16(a). Curve (1) represents the first term only, the others are assumed to be equal to unity. This may be called the ideal response, where no loss occurs and it has a level of 91.7 *db* vs. (1 dyne/cm^2)^2 per watt for a 6 cc. front volume and a matching frequency of 1000 cps. Curve (2) represents the product of the first two terms; added is the volume loss $V_c/V_g + V_c$, which is the effect of introducing additional front volume between the diaphragm and cap, thus increasing the total front chamber volume of the load. Curve (3) includes the effect of the third term of the equation, which may be called the stiffness loss, because it adds a diaphragm stiffness S_r, in series with the load stiffness, S_f. Curve (4) includes the coupling factor, $K/1 + K$, which depends on the force factor that can be developed in relation to the inductance and stiffness of the system. This factor contains the term $K = \dfrac{\tau^2 10^{-7}}{L(S_r + S_f)}$ which is a sort of coefficient of coupling, analogous to the electrical coupling coefficient of a transformer, except that it may exceed unity. The whole factor, however, $K/1+K$, can never exceed unity, and we may call it k^2, defining k as the "coupling factor." Thus each term of this equation may be associated with some physical part of the receiver, which contributes to the losses.

A fifth term will now be developed, which may be called the resistance loss, due to the electrical resistance of the receiver. If the receiver has a resistance R, the last term of equation (8) may be written as $\dfrac{R_0 X}{(R_0 + R)^2 + X^2}$ where R and X are taken as the measured low-frequency resistance and reactance of the receiver.

In equation (9) however, the term $\dfrac{f/f_0}{1 + (f/f_0)^2}$ was factored out of this expression and included as part of the first term. To take account of this, the remaining term for the resistance loss becomes

$$\text{Resistance Loss} = \frac{R_0 X}{(R_0 + R)^2 + X^2} \cdot \frac{1 + (f/f_0)^2}{f/f_0}$$

If this remaining factor is included, the expression for the response of the receiver with resistance becomes

$$10 \log \eta = 10 \log \left(\frac{2\gamma P_0 \, 10^7}{V_c \pi f_0 \left(1 + \frac{f^2}{f_0^2}\right)} \right) \cdot \left(\frac{V_c}{V_g + V_c} \right) \cdot \left(\frac{S_f}{S_r + S_f} \right)$$
$$\cdot \left(\frac{K}{1 + K} \right) \cdot \left(\frac{R_0 X}{(R_0 + R)^2 + X^2} \cdot \frac{1 + f^2/f_0^2}{f/f_0} \right) \qquad (10)$$

This is the equation of the curve (5) shown in Fig. 16a. This curve checks quite closely with the measured response of the actual receiver, shown by the dashed curve (5). The close coincidence of the solid and dashed curves constitutes a check on the accuracy of both the theoretical and measured response of the receiver. The slight divergence of these curves in the range from 300 to 500 cps is due to the effect of the mass of the diaphragm, which was neglected in the calculations.

While the above analysis is limited to low frequencies, it gives one an indication of the magnitudes of the various types of losses. It shows that, of these, the stiffness loss and coupling loss are the greatest. Considerable progress has been made in the design of this receiver in reducing the stiffness loss from 11.3 *db* for the HA1 receiver to 6.9 *db*. This is due largely to the increased effective area and low acoustic impedance of the diaphragm.

Receiver Efficiency Limits

In the analysis above, it is shown that for an ideal receiver having no losses, operating into a 6 cc. chamber, and matched at 1000 cps, the response approaches an upper limit of 91.7 *db* vs. (1 dyne per cm^2)^2 per watt. In other words, this is the limit which the low-frequency response of an idealized receiver approaches when the diaphragm stiffness S_r, the front chamber V_g, and the coil resistance R all approach zero, and the coupling factor k approaches unity. However, in addition to the level limit there exists a frequency range limit which lowers the level of the former limit. The curve labelled "Locus of Performance Limit" of Figure 16(b) determines both these boundaries. Thus, if any point is selected on this curve, a horizontal and a vertical line through it determine simultaneously the maximum response level and the highest frequency range obtainable. The calculation of both

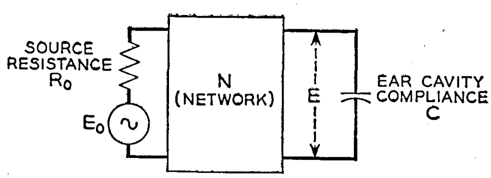


Fig. 17—Network of an ideal receiver having a uniform response over a given band of frequency.

limits may be based on H. W. Bode's resistance integral theorem.[13] In accordance with this theorem, when an ideal coupling network N, shown in Fig. 17, is used to give the maximum performance between a resistance source R_o and a capacitative load C, we have the general formula

$$\int_0^\infty e^{2a}\, d\omega = \frac{\pi}{2CR_0} \tag{11}$$

where $e^{2a} = \left(\dfrac{E}{E_0}\right)^2$

E_0 = the source voltage

and E = the voltage developed across the load C.

[13]"Network Analysis and Feedback Amplifier Design," H. W. Bode—D. Van Nostrand Co.—p. 362.

If a flat transmission is required over a given band of frequencies we may replace the limits of integration by ω_1 and ω_2, these quantities representing the edges of the useful band, which yields

$$\left(\frac{E}{E_0}\right)^2 \cdot \int_{\omega_1}^{\omega_2} d\omega = \frac{\pi}{2CR_0}$$

or

$$\frac{E^2}{E_0^2/4R_0} = \frac{2\pi}{C(\omega_2 - \omega_1)} = \frac{1}{C(f_2 - f_1)} \tag{12}$$

Translating this expression into the equivalent acoustical system, where C represents the compliance of the human ear cavity, it may be shown* that, by substituting $C = \dfrac{V_c}{\gamma P_o}$, replacing the voltage E by the pressure p developed in the cavity, and by using the factor of 10^7 to convert from practical to c.g.s. power units the following equation results

$$\eta = \frac{p^2}{E_0^2/4R_0} = \frac{\gamma P_0 .10^7}{V_c(f_2 - f_1)}. \tag{13}$$

This is the expression for the available power response of an ideal receiver which is assumed to have a flat response over the band of frequencies extending from f_1 to f_2.

The plot of this equation expressed in decibels, and marked "Locus of Performance Limit" is shown in Fig. 16(b) taking V_o as 6 cc and f_1 as zero. From this curve, an ideal receiver having a bandwidth of 3500 cps would have a response level of 88.3 *db*. It is also evident that the low frequency analysis given in the preceding section corresponds to a receiver with a range of approximately 1600 cps, while for wider ranges the response level would be lower, corresponding to the three other bands shown in the figure.

From the analysis given above, it is clear that the 88.3 *db* level limit supersedes the 91.7 *db* value based on the low frequency losses alone, because the former value takes account of the frequency range over which a receiver is designed to operate. A complete loss theory would undoubtedly arrive at the lower limit. However, because of the reactive load, it has not been possible to derive a suitable formula which includes the dependence on frequency range, and at the same time shows the character of the losses. The utility of the low frequency analysis lies in the fact that it shows the relative importance of the various losses, and where the most opportunity for improvements lies, and their likely magnitudes. It must be realized that although ring armature receivers may be built in the laboratory, which have smaller losses than the receiver discussed, the present design is a compromise chosen to be most suitable for use in the subscriber's telephone set.

*Unpublished work by T. J. Pope, Bell Telephone Laboratories, Inc.

Magnetic Circuit

The essentials of the magnetic structure of the ring armature receiver, including an equivalent circuit, are shown in Figs. 18(a) and 18(b). As explained earlier, the magnetic structure includes an L-sectioned ring pole-piece of 45% permalloy having an outwardly extending flange which carries a non-magnetic ring, the latter acting as a support for the permendur diaphragm. The remalloy magnet, which is also an L-sectioned ring, is assembled over the pole-piece assembly so that its inwardly extending flange overlies the diaphragm. This overlying portion of the magnet plays an important part in that it enhances the force factor of the device by securing some of the advantages of a balanced armature type of receiver in a simpler

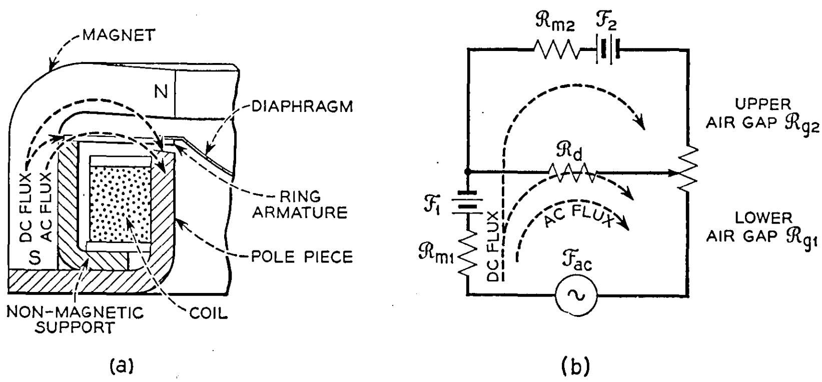


Fig. 18—Magnetic circuit analogy. (a) Physical arrangement of magnetic structure. (b) Equivalent magnetic circuit.

type of structure. The auxiliary magnet principle shown here was first used on the simple bipolar receiver,[5] and later was applied to the ring armature type structure.[4] In both cases, gains in force factor of 3 to 6 *db* were realized

The equivalent magnetic circuit of Fig. 18(b) shows to a first approximation the relations of the physical elements of Fig. 18(a), neglecting the leakage paths. As shown, the overlying portion of the magnet provides a shunt path so that a part of the d-c. flux flows around the armature and only through its inner marginal portion, while the lower portion of the magnet carries additional flux to the armature and through the main air gap. The magnetic circuit is a type of partially balanced circuit which, if fully balanced, will have no d-c. flux flowing through the armature provided the following relations are satisfied:

$$\frac{\mathcal{F}_1}{\mathcal{F}_2} = \frac{\mathcal{R}_{m1} + \mathcal{R}_{g1}}{\mathcal{R}_{m2} + \mathcal{R}_{g2}}$$

Where $\mathcal{F}_1$ = M.m.f. of lower cylindrical part of magnet
$\mathcal{F}_2$ = M.m.f. of flanged portion of magnet
$\mathcal{R}_{m1}$ = Reluctance of lower cylindrical part of magnet
$\mathcal{R}_{m2}$ = Reluctance of flanged portion of magnet
$\mathcal{R}_{g1}$ = Reluctance of main air gap (a variable modulating reluctance)
$\mathcal{R}_{g2}$ = Reluctance of auxiliary air gap
$\mathcal{R}_d$ = Diaphragm reluctance

The above relationship can be derived by placing the flux through the diaphragm reluctance $\mathcal{R}_d$ equal to zero in the circuit shown. Under these conditions, the armature carries no d-c. flux over its middle portion and will be operating at maximum permeability to a-c. flux. Moreover, the d-c. air gap flux density in the lower air gap where most of the field of force resides can be made higher before saturation begins to degrade the permeability of the inner marginal portion of the armature, than if all of the armature had to carry d-c. flux. The above factors tend to increase the a-c. and d-c. flux, and, since the force factor is a function of the product of these two quantities, a higher force factor will result from the addition of the overlying portion of the magnet.

In order to maintain the position of the freely supported diaphragm on its seat at the outer periphery, it has been found desirable to have only a partial balance of the circuit. This is accomplished by making the upper air gap approximately five times larger than the lower one. Thus the field in the upper air gap is weaker, so that a 25 to 50% unbalance in flux exists. Under these conditions, the flux component in the diaphragm due to the upper portion of the magnet only partially cancels that due to the lower portion. However, the resulting flux density in the diaphragm is such that the permeability will be only slightly below the maximum permeability which obtains for the perfectly balanced condition. The reluctance of the upper mesh to a-c. flux is so high, that the a-c. flux flowing in this branch can be neglected, hence the lower mesh carries substantially all of the a-c. flux, as shown in the figures. Thus, a partial separation of the a-c. and d-c. flux paths is accomplished.

The magnetic materials which comprise this structure include a remalloy magnet, a vanadium permendur diaphragm, and a 45% permalloy pole-piece. Some of the considerations which led to the choice of these materials are indicated below. The remalloy magnet can be formed from sheet material while at elevated temperatures, is machinable prior to the final heat treatment, and has good magnet properties. Although Alnico could be used as magnet material, it would not lend itself to forming, and the result would be a more expensive magnet. The vanadium permendur diaphragm has a higher permeability at the higher flux densities than other materials, and this results in a higher force factor. The high yield point and high modulus of elasticity of permendur give better elastic properties so that the diaphragm will restore over a wider range of deflections to which it may be subjected. The 45% permalloy pole-piece has a high resistivity, resists corrosion without the need of a finish to protect it, and is easily formed to the desired shape. Since it has a high permeability, and is not too sensitive to strains, it is well suited for pole-piece material.

Application of Ring Armature Receiver to Other Sound Devices

Several other applications of the ring armature transducer have been made on an experimental basis in addition to that of the handset receiver,

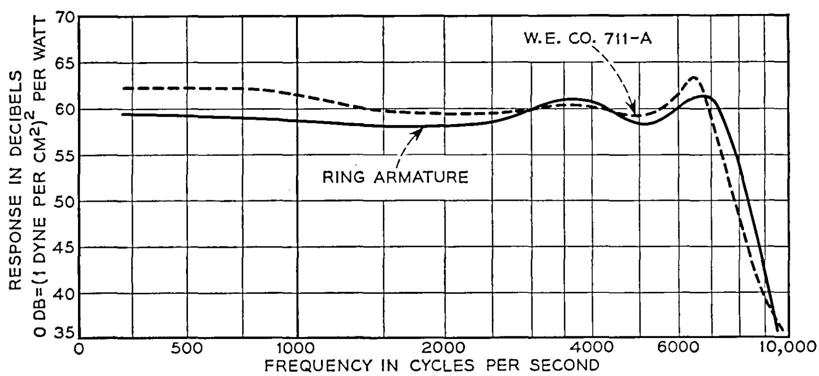


Fig. 19—Available power response-frequency characteristic of an experimental wide-range ring armature receiver, compared to a W.E.Co. 711-A moving coil receiver, measured on a 6 cc. closed coupler.

such as its use as a wide-range receiver, as a miniature horn-type loudspeaker, and as a microphone.

By narrowing the ring armature and suitably proportioning the acoustic networks, the receiver may be made to operate over a much wider frequency range at some sacrifice of efficiency, and may thus be used for high quality monitoring and audiometric work. A response characteristic of such a unit is shown in Fig. 19. The response compares favorably with the Western Electric 711-A moving coil type receiver used for the same purpose. Being comparable in efficiency, it has also a similar frequency range of about 7000 cps. More important, however, are the greater ruggedness and simplicity of the ring armature type receiver. Moreover, the impedance may be made to suit the application over a considerable range of impedance values, whereas the moving coil type is limited to a coil of about 25 ohms. While somewhat less pure in tone than the moving coil type the harmonics are 50 *db* below the fundamental at a sound pressure of 20 dynes per cm^2 in the coupler, and therefore are not noticeable under ordinary listening conditions. The attainment of wide frequency range in magnetic type units is unique, and is due in part to the use of a peripherally driven diaphragm. Centrally driven diaphragms used in magnetic type receivers usually have parasitic modes of vibration at these frequencies, which places a limit on the frequency range for which they can be designed.

As a loudspeaker, the ring armature structure has been found to have some experimental application when used with a horn, both for speech and as a sound source for measuring purposes. Response characteristics of such

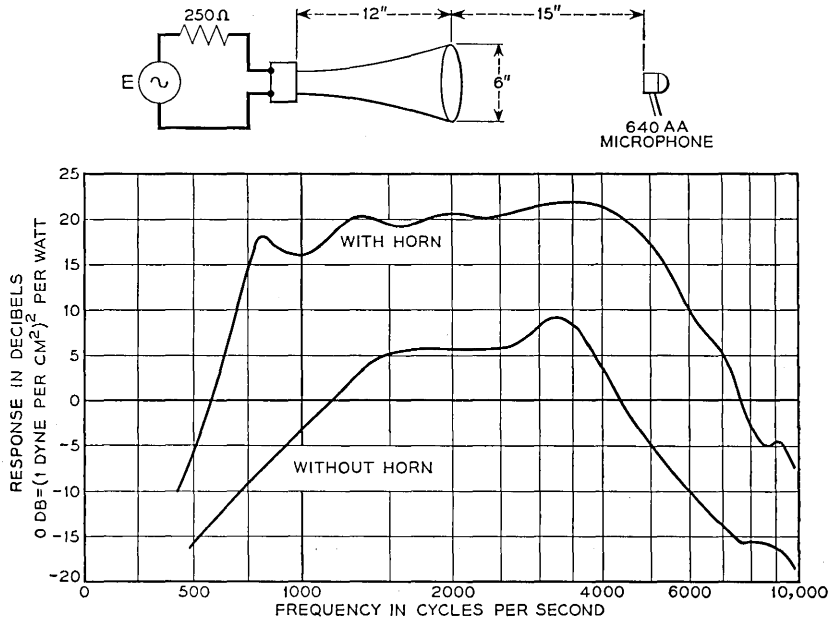


Fig. 20—Available power response-frequency characteristic of a ring armature receiver, measured as a loudspeaker with and without a horn.

a unit are shown in Fig. 20 for the instrument with and without a horn. In the case of the unit without a horn the unit had a receiver cap as used in the handset receiver, and the acoustic circuits were similar to those of the normal receiver. For the case of the horn attached to the unit a spherical plug was used to couple the horn closely with the diaphragm, and no acoustic damping circuits beneath the diaphragm were necessary. The horn pictured in the sketch was spaced 15 inches away from the measuring microphone, and the same distance was used for the curve without a horn. It is apparent that, with the horn, 15 db is added to the sound level on the axis, and the frequency is widened by a factor of 2 or more. The efficiency shown in Fig. 20

compares favorably with similar devices of the moving coil type. The maximum acoustic output, however, is lower, because the amplitude of the diaphragm is limited by the air-gap.

As a microphone, the ring armature structure may be modified to have characteristics which are quite favorable for certain types of applications. Figure 21 shows the field response of a ring armature unit modified for use as a microphone and measured on an open circuit voltage basis. A special housing of 30 cc. rear volume was used in this case, with a $\frac{1}{8}$ inch diameter orifice in the rear of the housing to act as a resonant circuit to produce the low frequency resonance shown with a desirable cut-off at 250 cps. By lowering the acoustic damping resistance of the unit, a second resonance was produced in the middle of the frequency range as shown. The peak at the upper end of the range is the normal characteristic of the instrument, but it may be enhanced somewhat by the use of a cavity in front of the diaphragm.

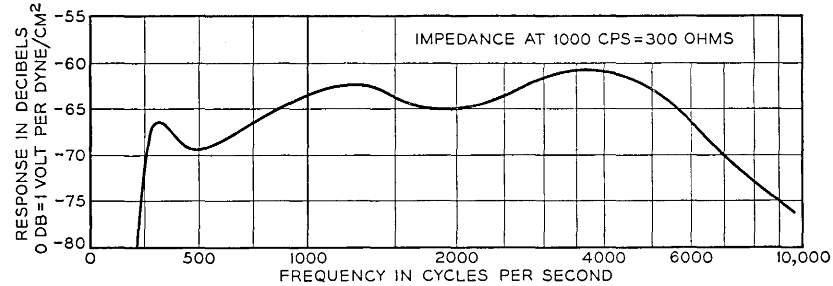


Fig. 21—Free field response-frequency characteristic of an experimental ring armature microphone, at normal sound incidence.

The output level of this microphone is about 9 *db* above the Western Electric 633A moving coil type at the same impedance level, but over a more limited frequency range.

Conclusions

It has been shown in the preceding sections that, by the use of the ring armature structure, it has been possible to realize a substantially higher grade of performance with regard to efficiency, frequency range, and leakage noise level, as compared to other types of telephone receivers in current use. To summarize, the ring armature receiver has been found to have the following advantages from a performance standpoint:

1. A gain in conversion efficiency of the order of 5 db as compared to the HA1 receiver and a corresponding increase in output capacity.
2. A wider frequency range, with an upper frequency of 3500 cps as compared to 3000 cps for the HA1 receiver.

3. A flexibility in frequency range, permitting the extension of the frequency range to approximately 7000 cps.
4. A broader superimposed direct current characteristic resulting in greater stability from a mechanical as well as an electrical standpoint.
5. A lower acoustic impedance, resulting in an improvement of the signal to ambient leakage noise ratio.
6. A substantial increase in the transmitted bandwidth when the receiver is held at small distances away from the ear.

Other advantages from a mechanical standpoint are:

1. A simple mechanical structure of ring-shaped or circular parts.
2. A low mechanical impedance, permitting the use of large air cavities and elements of low acoustic impedance for response control.
3. A concave, spherical dome-shaped diaphragm withstanding high transient pressures.

This work was carried on under the supervision of Messrs. W. C. Jones, F. F. Romanow, and W. L. Tuffnell, from whom many valuable suggestions were received. We also wish to acknowledge the valuable assistance of Messrs. P. Kuhn, L. A. Morrison, R. E. Polk, W. C. Buckland, R. R. Kreisel and R. E. Wirsching during the development of this receiver.

Internal Temperatures of Relay Windings

By R. L. PEEK, JR.

(*Manuscript Received Aug. 18, 1950*)

The steady state temperature distribution of a relay winding depends upon the power supplied and upon the rates of heat removal at the inner and outer surfaces. This is analyzed in terms of a more general form of the temperature distribution relation discussed by Emmerich (*Journal of Applied Physics*)[1]. This analysis is used to determine empirical constants for the rates of heat removal at the surfaces. Illustrative data are given for a stepping magnet.

Introduction

EMMERICH[1] has developed an expression for the steady state temperature distribution in a magnet coil when the heat flow is wholly radial, and the temperatures of the inner and outer surfaces are the same. A more general problem arises in the case of relays and other electromagnets used in telephone switching apparatus. In these cases, the coil is mounted on an iron core. Heat is withdrawn from the coil partly by conduction through this metal path, and partly by radiation from the outer surface. In consequence of this, the temperatures of the inner and outer surfaces are in general different.

In the relay and switching magnet problem, primary interest attaches to the rate at which heat is withdrawn through these two paths, as their combined effect determines the maximum temperature attained within the coil. The analysis outlined below has been employed to determine the division of heat between these two paths, and for the evaluation of empirical constants of heat removal. These constants are used in estimating the relation between the temperature of the winding and the power supplied to it.

Theory

As in Reference (1), it is assumed that there is no heat loss through the ends of the coil, so that the temperature gradient is wholly radial, and that the actually heterogeneous coil structure can be treated as homogeneous. Then if Q is the heat supplied per unit volume per unit time, and K is the thermal conductivity, the radial distribution of temperature is the solution to Poisson's equation:

$$\frac{d^2T}{dr^2} + \frac{1}{r}\frac{dT}{dr} + \frac{Q}{K} = 0, \tag{1}$$

[1] C. L. Emmerich, Steady-State Internal Temperature Rise in Magnet Coil Windings, *Journal of Applied Physics*, 21, 75, 1950.

where r is the co-ordinate of a surface of temperature T. The general solution to equation (1) is given by the equation:

$$T = A + B \log r - \frac{Qr^2}{4K}, \tag{2}$$

where A and B are constants, determined by the boundary conditions. The temperature has a maximum value T' at some radius r', at which the temperature gradient $dT/dr = 0$. Substituting the expression for T given by equation (2) in this condition, it is found that:

$$r'^2 = \frac{2KB}{Q}. \tag{3}$$

If the expression for B given by equation (3) is substituted in equation (2), and if A is taken as given by the resulting expression for T' when $r = r'$, equation (2) may be written in the form:

$$T = T' + \frac{Qr'^2}{4K}\left(1 + 2 \log \frac{r}{r'} - \left(\frac{r}{r'}\right)^2\right). \tag{4}$$

This equation gives the general expression for the temperature distribution in terms of the radius r' at which the temperature has its maximum value T'. In the special case in which the temperature T_1 at the inner radius r_1 is the same as that at the outer radius r_2, substitution in equation (4) of $r = r_1$ and $r = r_2$ gives two expressions for T_1. From these there can be obtained the same expression for the radius r' of maximum temperature as is given in Reference (1) for this special case. In the notation used here this expression is:

$$r'^2 = \frac{r_2^2 - r_1^2}{2 \log \dfrac{r_2}{r_1}}. \tag{5}$$

Substitution of this expression for r' in equation (4) gives an expression for $T - T_1$ which is identical with that given by equation (18) of Reference (1).

Using the expression for T given by equation (4), integration of $2\pi rTdr$ over the interval r_1 to r_2, and division of this integral by $\pi(r_2^2 - r_1^2)$, the coil volume per unit length, gives the following expresson for the mean coil temperature $\overline{T}$:

$$\overline{T} = T' + \frac{Qr'^2}{2K}\left(\frac{r_2^2 \log \dfrac{r_2}{r'} - r_1^2 \log \dfrac{r_1}{r'}}{r_2^2 - r_1^2} - \frac{r_1^2 + r_2^2}{4r'^2}\right). \tag{6}$$

EXPERIMENTAL

By means of equation (4), coil temperature measurements may be analyzed to determine the thermal conductivity and the rates of heat removal from the inner and outer surfaces of the coil. If the heat flow is wholly radial, all the heat supplied must pass through one or the other of these surfaces. The division of the heat between the two paths is determined by the radius r' of maximum temperature. The rate of heat flow to the core is therefore the rate of heat supply per unit volume Q, multiplied by the volume of the coil inside the radius r', or $\pi\ (r'^2 - r_1^2)$ per unit length of coil. Similarly, the rate of heat flow through the outer surface per unit length of coil is $Q.\pi\ (r_2^2 - r'^2)$.

It is therefore formally possible to determine the heat division by measuring the temperature distribution, and reading the radius of maximum temperature directly from it. When this is done it is found that the temperature gradient is comparatively flat in the vicinity of the maximum, and that it is therefore difficult to measure r' directly. An indirect determination of the radius r' may be made, however, by determining the maximum temperature T' and the temperatures T_1 and T_2 at radii r_1 and r_2 respectively. Expressions for T_1 and T_2 are obtained from equation (4) by letting $r = r_1$ in the one case and r_2 in the other. Then from these two expressions:

$$\frac{T_1 - T'}{T_2 - T'} = \frac{1 + 2 \log \frac{r_1}{r'} - \left(\frac{r_1}{r'}\right)^2}{1 + 2 \log \frac{r_2}{r'} - \left(\frac{r_2}{r'}\right)^2}. \tag{7}$$

Knowing T_1, T_2, T', and the radii r_1 and r_2, r' may be evaluated by numerical or graphical solution of equation (7). This solution is facilitated by the use of a table or plot of the function $Y\ (X) = 1 + \log X^2 - X^2$. The numerator of the right hand side of equation (7) is $Y\ (r_1/r')$, and the denominator is $Y\ (r_2/r')$.

A convenient procedure for determining the temperature distribution within a coil is to measure the resistance changes in different layers, tapping the coil between these layers. If there is more than one layer between taps, the resistance change measures the mean temperature of the layers included.

For an accurate determination of the gradient, it is convenient to make the resistance measurements on a comparative basis, as by use of the bridge circuit shown in Fig. 1. Here the terminals marked 0 and 6 are the inner and outer ends of the winding, while terminals 1 to 5 denote taps at intermediate layers. With the key K in Position 2, the bridge circuit may be balanced to establish the resistance between terminals 0 and 3, for example, relative to that between terminals 3 and 6. The resistance of the whole winding may be determined with sufficient accuracy (about one per cent) by voltmeter-ammeter readings made with K in Position 1. With this known, the resistance of the layers between taps, and hence the temperature differences, can be computed from the bridge readings.

The temperature corresponding to the resistance between taps may be taken as the temperature at the mean radius of the layers included. The

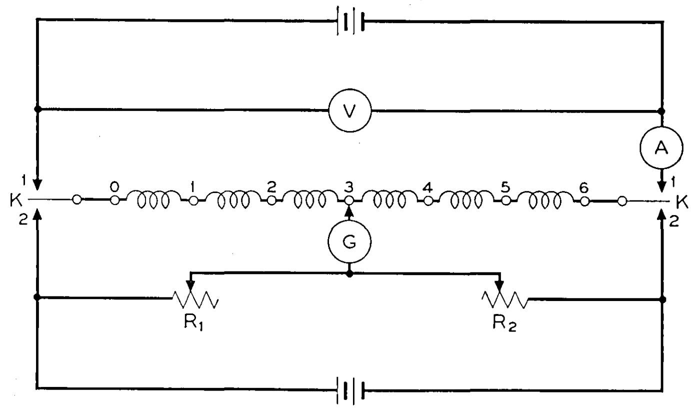


FIG. 1—Circuit for measuring temperature differences in tapped coils.

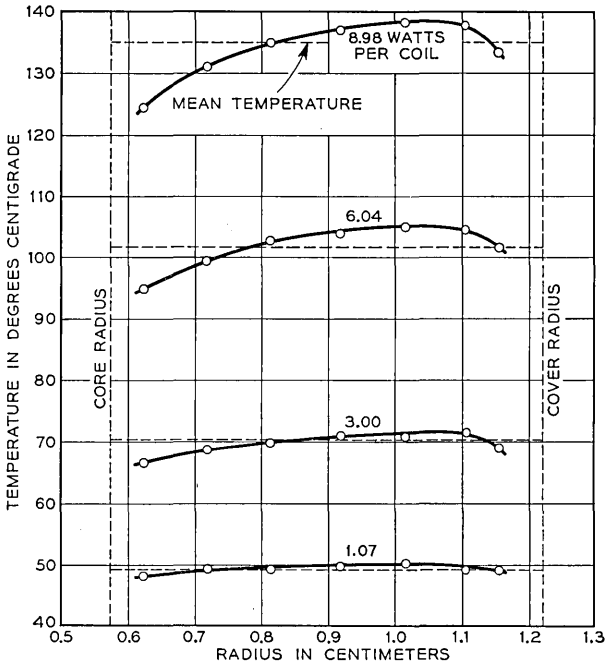


FIG. 2—Temperature distribution in coil of 197 switch magnet.

results of such measurements may be plotted as shown in Fig. 2. The results given in this figure were obtained with a sample of the Vertical Magnet of the Western Electric Company's 197 (Step-by-Step) switch.

The plotted points show the mean temperature of the layers between taps plotted against their mean radius; the dashed boundary lines show the inner and outer radii of the coil. The different curves correspond to the different levels of steady state power input indicated.

It is apparent in Fig. 2 that the central part of each curve is comparatively flat, and that it would therefore be difficult to determine the radius of maximum temperature directly. It may be noted in passing that the maximum temperature is not greatly in excess of the mean temperature. This justifies the usual engineering practice of taking the mean temperature as a criterion of whether the coil is overheated or not.

TABLE I
HEAT DISTRIBUTION IN 197 SWITCH COILS

Total Heat Input (Watts)	1.07	3.00	6.04	8.98
Q (watts per Cm.^{3})	0.096	0.268	0.540	0.802
Max. Temperature T' (°C)	50	72	105	139
Temperature T_1 (°C)	48	67	95	124
Temperature T_2 (°C)	49	69	102	132
K (Calories/°C/sec./cm.)	0.53×10^{-3}	0.69×10^{-3}	0.79×10^{-3}	0.82×10^{-3}
r' (Cm.)	0.92	0.93	0.95	0.94
Heat to Core (Watts)	0.48	1.39	2.99	4.37
Heat to Cover (Watts)	0.59	1.61	3.05	4.61
(Per Cent.)	55	54	51	51
Inner Surface Temperature (°C)	47	65	91	119
Outer Surface Temperature (°C)	49	68	99	130

From each curve in Fig. 2 there was read the maximum temperature T' and the temperatures T_1 and T_2 at $r_1 = 0.622$ cm. and $r_2 = 1.156$ cm. respectively, corresponding to the inner and outer points plotted. These temperatures are listed in Table I together with the corresponding values of Q, the power input divided by the volume of the coil (11.2 cm^3). From these data, the radius r' of maximum temperature has been computed by the procedure described above. With r' known, the quantity $\frac{Q}{K}$ was computed from equation (4) for the case $r = r_1$, $T = T_1$. The resulting values of r' and K are included in Table I. Using the value of r' thus determined, the division of the heat between that going to the core and that going to the cover was computed with the results shown in the table. The values of r_1 and r_2 used in the above computation are, as indicated in Fig. 1, internal to the coil. Taking new values of r_1 and r_2 corresponding to the core and cover radii respectively, the temperatures at these surfaces were computed from equation (4), using the values already found for r', T and $\frac{Q}{K}$. These core and cover temperatures are included in Table I.

It is of some interest to examine the observed values of apparent conductivity K. As the temperature differences for the two lower input measurements are small, the results for the other two cases are more accurate. The latter give a value for K of approximately 0.8×10^{-3} calories per °C per sec. per cm. In this form wound coil using No. 29 wire, the volume occupied by the insulation is approximately 36 per cent. As the conductivity of the copper is very large compared with that of the insulation, the conductivity of the latter may be estimated as of the order of K multiplied by the fraction of the volume occupied by the insulation. This gives an indicated conductivity for the insulation of the order of 3×10^{-4} calories per °C per sec. per cm. which is about the same as that of dry paper.

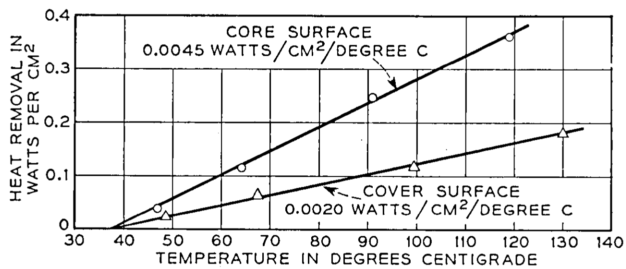


FIG. 3—Rates of heat removal from coil of 197 switch magnet.

For engineering purposes the results of major interest are those for the quantities of heat leaving the core and the cover. The results in Table I show that the heat flow is about equally divided between these two paths. As the core radius is about half the cover radius, the rate of heat flow to the core per unit area of surface is about twice that leaving unit area of the cover surface.

The rates of heat removal per unit area through the inner and cover surfaces are shown plotted against the corresponding temperatures in Fig. 3. It will be seen that the relation between the rate of heat removal per unit area and temperature is approximately linear, and intersects the temperature axis at 38°C (100°F), which was the ambient temperature in these tests.

It follows that for engineering purposes the rate of heat removal per unit area through either the inner or cover surface may be taken as proportional to the difference between the surface and ambient temperatures. A similar linear approximation has been found to apply for other relay and switch coils when mounted under conditions representative of telephone apparatus. Because of the multiplicity of mounting conditions and the complexity of conducting and radiating paths by which heat is removed, it is difficult to establish a relationship of the type shown in Fig. 3 by analysis of the paths of heat removal. For given mounting conditions, however, this relationship can be determined empirically by the procedure outlined above and used to estimate the heat removal from a given coil mounted under conditions for which such measurements have been made.

The rate of heat removal is thus measured by the slopes of the heat flow vs. temperature curves, which may be designated k_1 and k_2. Thus k_1 is the time rate of heat flow to the core per square centimeter of surface per °C difference between the inner coil surface and the ambient temperatures, while k_2 is the corresponding coefficient for the cover surface. For the case shown in Fig. 3, $k_1 = 0.0045$ watts per cm^2 per °C, and $k_2 = 0.0020$ watts per cm^2 per °C. This observed value of k_2 is characteristic of the cover surfaces of coils mounted as in telephone apparatus, where the heat removal is primarily by radiation to surfaces at or near the ambient temperature. While the value of k_1 observed for this case is representative of that applying to inner coil surfaces, the values of k_1 for such surfaces vary widely, and are particularly sensitive to variations in the clearance between the metallic core and the interior surface of the coil.

Prediction of Coil Temperatures

If values for the heat removal coefficients are known, the distribution of temperature within the coil for a given steady state power input may be determined from equation (4). The power input and the coil volume determine the rate of heat supply per unit volume Q. The rate of heat flow to the core per unit length of coil is therefore $\pi\ (r'^2 - r_1^2)\ Q$. The core area per unit length through which this heat passes is $2\ \pi r_1$. So from the empirical linear relation between the heat removed and the surface and ambient temperatures:

$$T_1 - T_0 = \frac{(r'^2 - r_1^2)Q}{2k_1 r_1}, \tag{8a}$$

where T_0 is the known ambient temperature. Similarly, for the cover surface:

$$T_2 - T_0 = \frac{(r_2^2 - r'^2)Q}{2k_2 r_2}. \tag{8b}$$

By substituting these expressions for T_1 and T_2 for T in equation (4), with r taken as r_1 in one case and r_2 in the other, there is obtained an expression for r' which reduces to the following equation:

$$r'^2 = \frac{\dfrac{r_1}{k_1} + \dfrac{r_2}{k_2} + \dfrac{r_2^2 - r_1^2}{2\mathrm{K}}}{\dfrac{1}{k_1 r_1} + \dfrac{1}{k_2 r_2} + \dfrac{1}{K} \log \dfrac{r_2}{r_1}}. \tag{9}$$

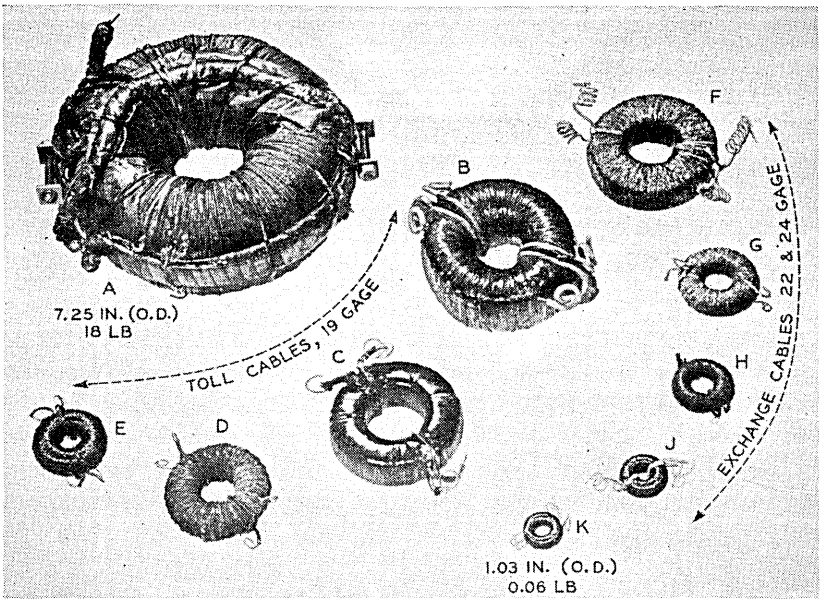
If r' is thus determined, and K is known from measurements of similar coils, T' can be determined by means of equation (4) from the values of T_1 or T_2 given by equations (8). If desired, the mean temperature $\bar{T}$ can then be determined by equation (6).

Conclusions

In most relay and switch magnet coils, of the type used in telephone apparatus, the heat flow under steady state conditions can be considered as wholly radial, and the temperature distribution conforms approximately to equation (4) above. In this expression, T' is the maximum temperature, and r' the radius at which this occurs, so that heat generated inside the surface of radius r' passes to the core, and that generated outside this radius passes to the cover. The rate of heat removal per unit area at either of these surfaces is found experimentally to be approximately proportional to the difference between the surface and ambient temperatures (for the temperature range of normal operation). The proportionality constant is the heat removal coefficient (k_1 or k_2 of equations (8)). Under conditions typical of telephone apparatus, this coefficient is of the order of 0.002 watts per cm^2 per °C for a cover surface, and 0.005 watts per cm^2 per °C for an inside surface in close proximity to the metal core. The heat removal coefficient for an inner surface is much more variable than that for a cover surface.

The coil temperature distribution [equation (4)] depends upon the rate of heat supply per unit volume Q, and upon the effective heat conductivity K. Q may be taken as equal to the total steady state power input divided by the coil volume. Correction might be made for the radial variation in Q resulting from the variation in copper resistivity with temperature. The relatively small temperature range observed in practice, as illustrated by the results of Fig. 2, makes this an unnecessary refinement. The heat conductivity constant K is an effective average value, applying to the coil as though it were a homogeneous structure. It is conveniently evaluated by measurements of actual coils, and is approximately a constant for a given wire size and type of insulation.

By measurement of the resistance changes in tapped coils, the internal temperature distribution can be determined. From this, values of the effective heat conductivity K and of the heat removal coefficients k_1 and k_2 can be determined by the use of equation (4), as described above. Conversely, if K and the heat removal coefficients are known, equation (4) may be used to estimate the internal temperature distribution, and thus the mean and maximum coil temperatures.



Headpiece.—A panorama of loading coils 1904–1948.

The Evolution of Inductive Loading for Bell System Telephone Facilities

By THOMAS SHAW

Introduction

THIS is the story of the contributions of inductive loading in the growth of the Bell Telephone System to its present great stature. In particular, it tells how these contributions have been made. We see in it the great economic value of organized research, of research scientists and design and development engineers, of manufacturing skill and care, and of the application of sound engineering principles to the design of the telephone plant.

The story told here is almost entirely concerned with series inductance coil loading, since other types of loading have had very little use in the Bell System. To make the story more complete, however, a brief account is included of Bell System developments and applications of continuous loading.

The main story starts soon after the turn of the century, following the development of the first standard loading coils and loading systems, and carries through to the end of 1949. In the beginning of this period emphasis was placed on urgently needed increases in the range of telephone transmission over open-wire lines and cables, and in the use of cheaper cables. In this connection it should be remembered that a really satisfactory type of telephone repeater did not become available for general use until 1915, about 15 years after the invention of loading.

The commercial limit in the economical use of loading to extend the transmission range of open-wire lines was reached in 1911, several years before the vacuum-tube repeater became available. Then, repeaters and loading had to be used as teammates to conquer the transcontinental distances. Subsequent improvements in repeaters, auxiliary equipment and circuits eventually made it advantageous to discontinue open-wire loading, and simultaneously opened an important new field for impedance-matching loading on the entrance and intermediate cables that unavoidably occur in open-wire lines. When different types of carrier systems became available for open-wire facilities, several new types of impedance-matching loading suitable for these carrier systems were developed.

The early efforts to extend the transmission range of long-distance cables in competition with open-wire lines, so as to obtain increased stability of service and lower facility costs, reached a climax during the period 1911–1915 in the use of composite, quadded, 10 ga. and 13 ga. cables, and of loading coils nearly as large as the open-wire loading coils. This trend slowed down shortly after vacuum-tube repeaters became commercially available. During the next fifteen years or so, intensive development work on improved repeaters, on equalizing and regulating networks, and on higher velocity, higher cut-off loading, made it feasible to use 19-gauge conductors and loading coils no larger than the initial standard cable coils for distances ranging up to about 1500 miles.

In the exchange area cable plant, coil loading has made it possible at a low cost to meet the needs imposed by geographical factors, with as yet very little competition by telephone repeaters. Large reductions in the costs of the trunk plant have resulted from the extensive utilization of 22- and 24-gauge cables, made feasible by the use of inexpensive loading. The substantially continuous transmission developments in exchange area services also made possible important improvements in the intelligibility of transmission by using higher cut-off loading to transmit wider speech-frequency bands.

The important loading apparatus developments in the period covered by the review have taken full advantage of the development at fairly even-spaced intervals of a series of successively better magnetic core-materials to improve the transmission service performance or reduce loading costs, sometimes combining these features. The loading coil cost-reductions which resulted from the large size-reductions made possible by the standardization of compressed permalloy-powder core loading coils during the late 1920's were especially important in influencing the growth of the long distance and exchange area cable plant, and in leading to important service improvements.

As described, step by step, in the present story, coil loading has been a very important factor in making possible the provision of satisfactory telephone service at reasonable rates which have encouraged a continually increasing use. Important elements in the public satisfaction to which loading has made fundamental contributions are: (1) high-quality transmission, and (2) high-speed service facilitated by the provision of relatively large groups of relatively low-cost facilities. In the extensive utilization of the long-distance service over repeatered, loaded, voice-frequency toll cable facilities, loading must of course share the credit for the improved transmission, plant cost-reduction, and speed of service with the telephone repeaters and associated equalizing and regulating networks, where involved.

All of the coil loading development work for Bell System needs, including the specific developments described in the present review, has been done by Bell System people without outside aid. Coil loading was independently invented by Dr. G. A. Campbell[1] of the headquarters staff of the American Bell Telephone Company, and by Professor M. I. Pupin[2] of Columbia University, at nearly the same time, in 1899. The patent interference proceedings made necessary by the conflicting claims of the Pupin and Campbell applications resulted in a priority award to Pupin during April 1904, on the basis of a few days' earlier disclosure. The prompt purchase of Pupin's rights in the invention before the interference action had gone far assured the Telephone Company complete freedom to develop the new loading art in the most advantageous ways.

The improvements worked out and applied over the years are principally due to groups of scientists and engineers working as teams on various phases of the transmission research, development, and engineering problems; on the magnetic materials research and development problems; on the apparatus-design and manufacturing problems; and on the field-construction and traffic problems. Nearly all aspects of telephone systems' development have been involved to a greater or less extent.

In the aggregate, a large number of individuals have made important contributions to the advancement of the loading art. The writer of this review is to be regarded as a spokesman for his co-workers. Since the assignment of a fair measure of personal credit to each individual who has been involved would be extremely difficult, it is not attempted in the present review.

PART I: THE BEGINNINGS OF COIL LOADING

General Theory

For present purposes, a rigorous presentation of the mathematical theory of coil loading is unnecessary.[a] A simple description and a brief statement of theory is sufficient.

The primary purpose of coil loading is to improve the transmission of intelligence by substantially reducing the circuit attenuation, and by making the circuit attenuation approximately uniform throughout a predetermined frequency-band. These transmission benefits are obtained by serially inserting coils having uniform inductance values at regularly recurring intervals along the circuit, but are limited to a frequency-band below the loading cut-off frequency. This is an inverse function of the square root of the product of the coil inductance and of the mutual capacitance of the loading sections between successive coils, as determined by the coil spacing and unit-length capacitance of the circuit. Above the loading cut-off frequency, there is a substantial suppression of transmission.

For more than a decade prior to Campbell's and Pupin's 1899 researches, the theoretical possibility of improving transmission over telephone lines by artificially increasing their inductance had become known from the mathematical studies of Vaschy and Heaviside. Also there had been considerable speculation by them[4, 5], and by others, regarding the practicability of approximating the advantages of uniformly distributed inductance by inserting low-resistance inductance coils along the line. Rules for spacing the lumped inductances had not been worked out, however, nor had suitable coils been developed.

The requisite coil-spacing turned out to be such that there are several coils per wave length at the highest frequency which should be efficiently transmitted to obtain satisfactory intelligibility. Here, "several" means more than two, since at the theoretical cut-off frequency there are two coils per wave length. In terms of the nominal velocity of propagation of the "corresponding smooth line" (a hypothetical line having the same total inductance and capacitance) there are π coils per wave length at the cut-off frequency; expressed in "loads-per-second," this nominal velocity is exactly π times the cut-off frequency in cycles per second.

The attenuation improvement obtainable with loading corresponds somewhat to the increase in impedance that results from the increase in induct-

[a] Readers interested in the rigorous mathematical theory are referred to Bibliography items (1) and (2). Campbell's treatment has been extensively used by communication engineers because of its comprehensive coverage of the frequency band concept in which the cut-off effects on propagation and impedance are emphasized. Also, his disclosures include explicitly the effects of conductor resistance and ratio of coil resistance to conductor resistance. His general formulas include the distributed inductance and leakage.

ance. This can be understood from the fact that in the (low-impedance) non-loaded circuit, the series dissipation losses which are proportional to the square of the line current are ordinarily very large relative to the dielectric dissipation (i.e. shunt) losses which are proportional to the square of the line potential. When the line impedance is increased a suitable amount by the loading, the decrease in series losses is much greater than the increase in shunt losses. In commercial practice, economic considerations generally prevent the use of high loading impedances which would result in the shunt losses becoming as great or greater than the series losses.

In situations where voice frequency attenuation improvement is the principal objective, the unit transmission loss can usually be reduced to the order of one-third to one-fourth of the non-loaded value. The loss reduction is less than this at low voice frequencies and more at high frequencies, resulting in a much more uniform transmission of the important frequencies that are required for intelligibility and naturalness. In certain situations which will be discussed later a lower ratio of attenuation reduction is accepted in order to obtain other, more important, transmission advantages.

Pioneering Developments

General

A full account of the pioneering research and development work would take much more space than is available in a review devoted primarily to the evolution of the loading art. The present account is therefore limited to a brief description of the first loading systems and apparatus standards that resulted from the pioneering work.[b]

Although the success of the 1899 laboratory investigations, and the Bell System's 1900 experimental installations on exchange cables and on open-wire lines, quickly built up a substantial demand for loading, the commercial applications had to be deferred pending the development of satisfactory types of loading coils. Then there followed a series of what should be considered as trial installations of different types of loading, tailored to the specific needs of particular projects. Analyses of the performance characteristics of these installations, supplemented by continuing experimental work in the laboratory and by engineering cost-studies, resulted in the establishment of a series of standard cable loading systems for general use late in 1904. The commercial development of satisfactory open-wire loading encountered many even more complicated problems than those involved in cable loading, and in consequence the standard loading for 104-mil lines did not evolve until 1905. This same type of loading be-

[b] A more complete description of these standards is given in Bibliography Reference (6). Reference (7) is also of interest. Reference (1) gives some details of Campbell's early work.

came standard for 165-mil lines in 1910, after a long period of additional development work to get better line-insulation.

All of this early loading development work was for non-quadded cables and for non-phantomed open-wire lines.

Loading Coils

The earliest speculative suggestions regarding coil loading recognized the critical need for obtaining a low ratio of coil resistance to circuit resistance, and by implication a low ratio of coil resistance to coil inductance. As was expected, this turned out to be a difficult design and manufacturing problem, especially with open-wire loading which was given development priority.

By April 1901, a very satisfactory coil-design had been worked out for open-wire loading by Mr. H. S. Warren, an associate of Dr. G. A. Campbell. It had a toroidal core, formed by winding a bundle of insulated mild-steel wires, 4 mils in diameter, on a suitably shaped spool, several miles of wire being used in each core. The manufacturing process of the outside supplier

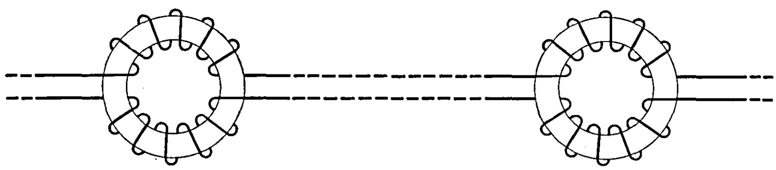


Fig. 1—Non-phantom type loading coils. Coil winding schematic and method of connection into circuit.

included cold drawing to obtain a magnetically hard wire having an initial permeability of about 65. The two line-windings, each confined to separate halves of the core winding-space, made use of insulated, stranded wire. The fine subdivision of the magnetic material and of the copper conductor was essential to the satisfactory control of eddy current losses.

This coil, Code No. 501, was the first standard loading coil. It remained standard for about a decade, until a redesign became necessary to facilitate an extensive commercial exploitation of phantom working. Because of the very low-resistance design objective, it had to be a large coil. Some of its dimensional and electrical characteristics are included in Table I. In size, it was approximately 25% larger than the largest coil shown in the head-piece.

The cable loading coils listed in the table were standardized for general use during 1904, following occasional use of other types of coils having different inductances and in some instances using a different core material. These coils were generally similar in their basic design features to the open-wire loading coil, but for economic reasons were much smaller in size.[c]

[c] Core weight 3.5 lbs., 69000 turns iron wire; length 11 miles.

Also, their ratios of resistance to inductance were considerably higher. In size, these coils are typified by Coil B in the headpiece.

The 4-mil (No. 38 A.W.G.) wire used in these cable coil cores had a nominal initial permeability of 95, nearly 50% higher than that of the open-wire loading coil core-material. The differences in magnetic performance characteristics, including permeability, core losses, and certain other features subsequently discussed, resulted solely from differences in the annealing treatments during the wire-drawing process.

Additional important differences between the cable coils and the open-wire coil were: (1) the use of non-stranded conductors in the windings of the cable coils, and (2) the use of lower dielectric-strength insulation in the cable coils.

TABLE I
CHARACTERISTICS OF FIRST STANDARD LOADING COILS

Code No.	Nominal Inductance (henry)	Resistance (ohms)		Use	Over-all Dimensions (inches)	
		d.c.	1000 cycles		Diameter	Axial Height
501	0.265	2.5	5.9	Open-Wire Lines	9	4
506	0.250	6.4	22.3	Cables	4½	2¼
508	0.175	4.2	13.0	"	4½	2¼
507	0.135	3.2	9.1	"	4½	2¼

Loading Coil Cases

The open-wire loading coils were individually potted in cast-iron cases designed for mounting on pole fixtures. The coil terminals issued from the case in individual rubber-insulated leads.

The cable loading coils were assembled in multi-coil groups on wooden spindles, and potted in cast-iron cases, suitable for installation in cable manholes and on pole fixtures. Intercoil crosstalk was controlled: (1) by mounting adjacent coils so there would be approximately a minimum coupling between their (small) magnetic-leakage fields; (2) by using iron shielding-washers between adjacent coils; and (3) by placing the spindle groups of coils in individual compartments cast in the cases. The coil-terminal leads issued from the cases in a twisted-pair, lead-covered, stub cable.

Prior to potting, the coils were thoroughly dried out under vacuum, and were impregnated with moisture-proofing compound.

Loading Systems

As previously indicated, the standard practices for cable loading were established in advance of those for open-wire loading, notwithstanding an earlier commercial beginning.[d] Three different cable loading standards were adopted, to provide for a range of attenuation-reduction performance. Some data regarding these systems are given in Table II. The transmission data apply to the early types of cable having an average mutual capacitance of about 0.070 mf/mi.

The theoretical loading cut-off frequencies were approximately 2300 cycles (about 7000 loads per second). This initial standard was the result of extensive series of speech transmission tests to determine the minimum cut-off frequency that would be commercially satisfactory with respect to intelligibility. A materially higher cut-off would have increased the loading costs by requiring the loading coils to be more closely spaced.

TABLE II
FIRST STANDARD CABLE LOADING SYSTEMS
(USING COILS OF TABLE I)

Loading Designation	Coil Inductance (henry)	Coil Spacing (miles)	Nominal Impedance (ohms)	Nominal Velocity (mi/sec.)	Attenuation Loss (db/mile)		
					19 A.W.G.	16 A.W.G.	13 A.W.G.
Heavy	0.250	1.25	1800	8750	0.28	0.16	0.11
Medium	0.175	1.75	1300	12200	0.39	0.21	0.14
Light	0.135	2.5	900	17500	0.51	0.27	0.17
Non-Loaded Cable	....	....	....	....	1.05	0.74	0.59

Note: The figures given in the columns headed "nominal impedance" and "nominal velocity" apply for the nominal impedances and the nominal velocities of the hypothetical "corresponding smooth lines," having the same total inductance and total capacitance.

The first standard open-wire loading used No. 501 coils at about 8-mile spacing, giving an impedance of about 2100 ohms and a cut-off frequency close to the standard cut-off frequency for cable loading. Under dry-weather insulation conditions (5 megohm-miles or better) the attenuation losses in the 104-mil and 165-mil lines were about 0.031 and 0.014 db/mi, respectively. The corresponding losses without loading were 0.075 and 0.033 db/mi, respectively. The approximate 8-mile spacing fitted in with the open-wire transposition arrangements and gave a satisfactory attenuation-loss reduction. The earlier attempts to secure a much greater attenuation reduction had involved shorter spacings, ranging down to 2.5 miles, and were unsuccessful. At extended periods of low line-insulation caused by wet weather, these higher-impedance loading arrangements had poor transmission, sometimes worse than non-loaded lines. Excessive noise, crosstalk, and reflection losses also were unfavorable factors.

[d] N. Y.-Chicago, 165-mil open-wire line, November 1901; New York-Newark cable, August, 1902.

Preview of Subsequent Developments

General Outline

For convenience in discussion and ease of understanding it has been found desirable to divide the remaining subject matter of this review into several parts, each covering a particular phase of the evolution of the loading art, as follows:

Part II—Loading for Long Distance Circuits.
III—Loading for Exchange Area Cables.
IV—Cable Loading Coil Cases.
V—Loading for Incidental Cables in Open-Wire Lines.
VI—Continuous Loading.
VII—Extent of Use and Economic Significance.
VIII—Summary and Conclusion.

Parts II, III, IV, V, and VII are wholly concerned with coil-loading.

In Parts II, III and V, specific coil loading systems and loading apparatus developments are separately considered under headings which indicate the development emphasis. In general, the individual developments are discussed in chronological sequence so as to tie closely together the interrelated systems and apparatus developments. The chronological procedure also applies in Parts IV and VI. The dates which are given and the cross references from section to section permit the reader to fit the important developments into a definite time pattern.

Although the review is primarily concerned with the evolution of loading, and its contributions to the growth of the Bell System, appropriate references are also included regarding other related advances in the telephone art which have influenced the design and performance of the loading systems and apparatus, and the extent of use.

Loading Systems

The changes in voice-frequency loading systems have been primarily for the purpose of improving the service performance, including the transmission of wider frequency-bands to improve intelligibility. The loading changes for repeatered circuits have catered to the various special problems that arose in consequence of the great increase in circuit lengths.

The loading systems for cable circuits transmitting radio broadcasting programs and those for voice-frequency and carrier-frequency impedance-matching in incidental cables that occur in open-wire lines also had their own individual requirements to meet the specific service needs.

Loading Coils

The loading coil developments substantially paralleled the loading systems developments in variety and scope. In many instances, new loading coils were developed to take advantage of the availability of (new) superior materials, improved design techniques, and fabrication methods, for cost reduction or service improvements. In other important instances, new types of loading coils were necessary to provide for new types of facilities; for example, (1) to permit the commercial exploitation of phantom working, (2) for a network of coarse-gauge, long-distance cables, (3) for facilities

TABLE III
LOADING COIL CORE-MATERIALS

Item No.	Type of Material	Effective[(a)] Volume Permeability	Approx. Period Commercial Mfg.	Principal Fields of Use	Bibliography References —Prior Publications
(1)	65-permeability, 4-mil, iron wire.	36	1901–1924 1911–1927	Open-Wire Lines Toll Cables	(6) & (8)
(2)	95-permeability, 4-mil, iron wire.	52	1904–1911 1904–1916	Early Toll Cables Exchange Cables	(6) & (8)
(3)	Annealed, compressed, powdered iron.	55	1916–1927 1916–1924	Exchange Cables Toll Cables	(6), (8) & (13)
(4)	Unannealed, compressed, powdered iron.	35	1918–1928 1924–1928	Toll Cables Exchange Cables	(6), (8) & (13)
(5)	Compressed, powdered permalloy.	75	1927–1937 1927–1938	Exchange Cables Toll Cables	(24)
(6)	Compressed, powdered molybdenum - permalloy.	125	1937– 1938–	Exchange Cables Toll Cables	(26)
(7)	Compressed, powdered molybdenum - permalloy.	60	1948–	15 kc Cable Program Transmission	(26)
(8)	Non-magnetic.	1	1920–	Carrier Loading Coils for Incidental Cables Open-Wire Lines	(8)

(a) Initial permeability.

to transmit programs for radio broadcasting stations, and (4) for incidental cables in open-wire carrier systems.

Certain economic concepts have dominated the design work on the individual loading coils. In the introductory section of the review, the general need for having the loading coil resistance low relative to that of the circuit resistance was mentioned. In applying this design rule, the principle of cost-equilibrium[(e)] has been a basic criterion. It has resulted in the de-

[(e)] This is a condition of cost-balance in which a small transmission improvement can be made by improving the coils at about the same cost as would be involved in improving the circuit in other ways—for example, by using a slightly larger size of conductor.

velopment of: (1) large-size, very low resistance coils for open-wire lines and for 10 ga. and 13 ga. long-distance toll cables, (2) smaller-size, higher-resistance coils for smaller-gauge toll cables, and (3) still smaller size of coils having still higher resistances for fine-wire exchange cables. Over the years, the progressive use of superior core-materials has made possible several successive, substantial size-reductions in the cable coils, with somewhat larger-ratio size-reductions in exchange area loading coils than in the toll cable coils, because of their less complex service-requirements and also in conformity with cost-equilibrium criteria. These progressive size-reductions are well illustrated in the headpiece.

Improved magnetic materials have been very important factors in the loading coil development work. The different magnetic materials which have been used in standard loading coils are listed in Table III with approximate dates, in terms of the beginning and end of manufacture, and other pertinent data.

PART II: LOADING FOR LONG-DISTANCE CIRCUITS

The early applications of standard open-wire loading made loaded 104-mil circuits about as good from the attenuation standpoint as non-loaded 165-mil circuits of equal length. When this loading was later applied to 165-mil circuits, the first New York-Denver loaded 165-mil line (1911) was approximately equivalent in transmission performance to the original non-loaded New York-Chicago 165-mil line (1892).

Large economies also resulted from the application of the first standard cable loading to suburban trunk cables and toll connecting trunks in exchange areas, and to interurban toll cables. Notable examples of the latter were the Boston-Worcester (1904), New York-Philadelphia (1906), and New York-New Haven (1906) cables. The toll cables used heavy loading. Considerable medium loading was used in long exchange cables.

(1) Phantom Group Loading

This was the first major new loading development to follow the pioneering standardization work. Beginning late in 1907, it culminated in commercial applications on open-wire lines and on new quadded cables during 1910.[f]

Entirely new types of coils were developed for loading the phantom circuits. Each of its four line-windings comprised a tandem connection of an inner-section winding located on one core-quadrant and an outer-section winding located on the opposite core-quadrant, the two line windings associated with the same side circuit being distributed over the same pair

[f] A much more comprehensive story of the phantom loading development and its relation to the development of quadded cable and to phantom working on open-wire lines is given in Bibliography items (8) and (9).

of opposite core-quadrants. The line windings were inserted in the four line wires of the phantom circuit and connected to have all of the mutual inductances aid the self inductances. In the individual side circuits the mutual inductances opposed the self inductances, and in consequence the coils contributed negligible leakage inductances to the circuits. The phantom coils were about twice as large in weight and volume as the associated side circuit coils, for cost-equilibrium reasons.

The side circuit coils were closely similar in size to the existing standard non-phantom coils. Each of their two line windings consisted of an inner-section winding on one half-core and an outer-section winding on the other

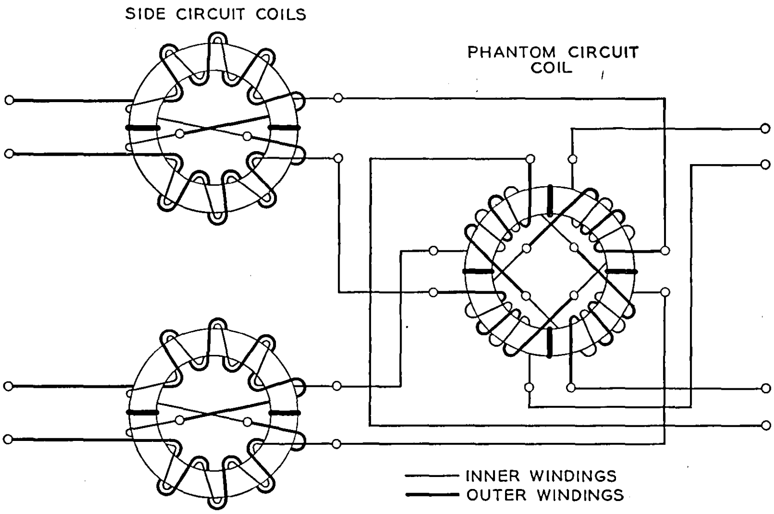


Fig. 2—Phantom group loading. Coil winding schematics and method of connection into circuit.

half-core, and thus in effect were evenly distributed about the entire core. The close magnetic-coupling thus obtained resulted in a negligible leakage inductance to the phantom circuit in the parallel-opposing connections of line windings. The mutual inductances aided the self-inductances in the side circuits.

Figure 2 schematically illustrates the coil winding arrangements. The general design symmetry of the individual coils also included essential symmetry in the distribution of the direct admittances among the line windings and from the line windings to the core and the case. The initial designs so well satisfied the service needs that only a very few minor design refinements were subsequently required from the crosstalk standpoint. The real difficulties encountered in meeting the service crosstalk-requirements were in controlling or correcting the small accidental unbalances that were unavoidable in manufacture.

The transmission performance in loaded side circuits was about the same as that of loaded non-phantom circuits on similar-size conductors. A slight attenuation impairment resulted from the non-inductive resistance of the phantom coils.

The phantom coils were located at side circuit loading points and the phantom inductance was chosen to give a cut-off frequency of about 2300 cycles, the same as in the side circuits, and in non-phantomed circuits. In consequence, the nominal impedance of the loaded phantoms was approximately 60% of that of the associated side circuits. The attenuation

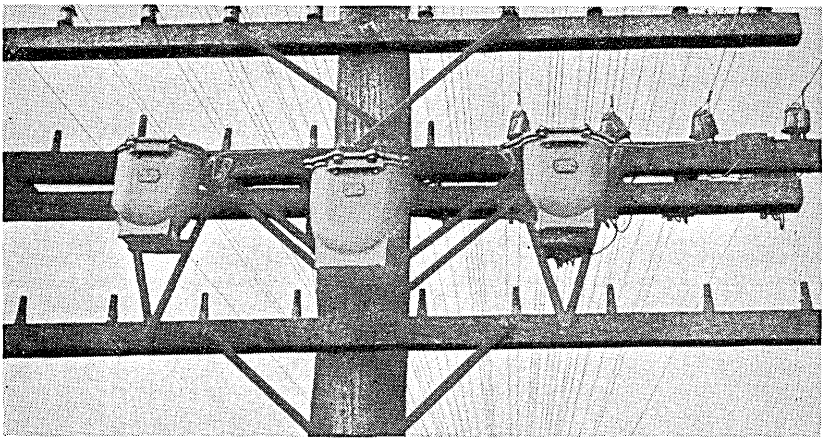


Fig. 3—An early installation of open-wire phantom group loading. Individually potted coils; phantom coil on pole; side-circuit coils on crossarms.

was about 13% better than that of the associated side circuits of open-wire lines, and from 15 to 20% better in loaded cables, depending upon conductor size.

The great commercial importance of the phantom-group loading development is indicated by the fact that nearly two-thirds of all the voice-frequency loading coils installed on quadded toll and toll entrance cables are coils of side circuit type, and nearly one-third are phantom loading coils. Over the years during which phantom loading and quadded cable have been available, only a relatively small amount of non-phantom type loading has been used in voice-frequency toll cable facilities. Important facilities in this special category are the loaded cable program-transmission circuits subsequently described and the "order wire" maintenance circuits in coaxial cables. Also, during the 1940's, there was some occasional use of non-phantom type coils on non-quadded exchange type cables for short haul toll facilities, in place of phantom group loading on quadded toll cables.

(2) Loaded Coarse-Gauge Quadded Toll Cables

The need for extending the telephone transmission range in storm-proof toll cable, and the unavailability of telephone repeaters suitable for use on loaded circuits, led to a great activity in the development of composite coarse-gauge quadded cables and of entirely new loading coils for these cables, beginning during 1910. The first application (1911) was the new Philadelphia-Washington Section of the Boston-Washington underground cable system.[g]

The new coils for the 10 AWG sides and phantoms were generally similar in design to the open-wire loading coils, even in their use of stranded copper conductors, and were only about 20% smaller. The coils for the 13-gauge conductors were intermediate in size (approx. geometrical mean) between the coils for the 10-gauge conductors, and the coils previously developed for 19 and 16 ga. cables. The size and efficiency relations among these three series of coils were approximately in cost-equilibrium for the grades of cable involved.

In new underground sections of these coarse-gauge cables a new standard "medium-heavy" weight of loading was used, the coil spacing and inductance being midway between those for standard heavy and medium loading and having the same cut-off frequency. The new weight of loading was nearly as good in transmission performance as the heavy loading that was used in old parts of the Boston-Washington route and other routes where the loading vaults had been laid out for heavy loading, and was considerably less expensive. In the medium-heavy loaded 10-gauge circuits, the attenuation was 0.050 and 0.040 db/mi, respectively for the sides and phantoms; the corresponding values in the 13 ga. circuits were 0.069 and 0.085 db/mi, respectively. The 10 ga. loaded circuits were designed for service between Boston and New York, and between New York and Washington. On an emergency basis, the phantoms could be used for Boston-Washington service.

It is appropriate at this point to mention the substantial reduction in toll cable dielectric losses that was worked out in the period under discussion. The extensive use of loading for the first time on long 10 gauge and 13 gauge circuits greatly increased the importance of reducing the amount of moisture that unavoidably accumulated in the conductor insulation during

[g] A more comprehensive account of this development and associated quadded cable developments is given in References (8) and (9).

the early stages of cable manufacture. This was done by refinements in the cable drying treatments.

(3) Changing Fields of Use for Iron-Wire Core Loading Coils

The new coarse-gauge cable loading coils, above referred to, marked the beginning of the use of 65-permeability iron-wire in place of 95-permeability wire in the cores of standard cable loading coils.

In every respect except permeability, the 65-permeability wire was superior to the higher permeability wire. The lower permeability was relatively disadvantageous as regards d-c resistance per unit inductance in coils of a given size. On the other hand, the core-loss resistance was substantially smaller, by virtue of the lower permeability and the superior hysteresis characteristics. In consequence, the total effective resistance of the 65-permeability core coils was lower at the upper speech-frequencies and nearly the same at the important middle frequencies, so that there was considerably less attenuation-frequency distortion.

Other even more important service advantages of the 65-permeability core toll cable loading coils resulted from their much greater magnetic stability. D-c signaling currents caused smaller temporary changes in inductance and effective resistance, in consequence of the superimposed d-c magnetization. Also, the residual effects of strong superimposed currents, manifested as permanent or semipermanent changes in inductance and effective resistance, were much smaller.

A specially valuable advantage of the 65-permeability wire core-material was in the substantially smaller amount of telephone transmission distortion caused by the operation of superposed composite telegraph systems.[10] The transient core-magnetization caused by the telegraph currents caused small transient changes in the inductances of the coils, and relatively very large transient changes in the effective resistances. The resulting non-linear distortion became known as "telegraph flutter." It varied as a function of telephone frequency and telegraph speed, the size of the core, the inductance of the windings, and the ratio of the amplitudes of the telephone and telegraph currents. It was accumulative in effect as the circuit lengths increased. Since simultaneous telephony and telegraphy was very general and was important from the revenue standpoint in the open-wire and cable long-distance facilities, the control of "telegraph flutter" became an increasingly important requirement in the development of new loading coils.

(The need for satisfactory control of "telegraph flutter" eventually led to the development of the improved cable telegraph systems which are described in Section 8 of this review.)

By 1912, the use of 95-permeability core-material in new toll cable coils had stopped. However, since exchange area circuits and suburban trunk cables were seldom used for composite telegraph working, the use of 95-permeability iron-wire continued standard until 1916, when compressed, powdered-iron, core coils became available.

(4) Loaded Repeatered Open-Wire Lines

The pioneering phases of the development of better lines and better loading coils for use on repeatered long-distance facilities had their first commercial application in the transcontinental open-wire circuits between New York and San Francisco, January 1915. The adaptation of the lines to the requirements of repeater operation was secondary in importance only to the development of satisfactory repeater elements, and of circuits for associating the repeater elements with the line. A comprehensive account of all phases of the very important transcontinental telephony-development project has been published in an article, "The Conquest of Distance by Wire Telephony."[9] Accordingly, the account in the present review is limited to the loading for the line. Comprehensive information regarding telephone repeaters is given in a 1919 paper by B. Gherardi and F. B. Jewett.[11]

Since the lines were used for two-way transmission, a high degree of impedance balance between the line and the repeater balancing-network circuit was necessary in order to obtain satisfactory repeater gains. This problem involved the construction of lines having a new order of regularity and stability in their impedance characteristics over the working frequency-band, to make feasible the design of simple types of balancing networks[12] for adequate simulation of the line impedances.

The requirements just stated involved a much greater degree of uniformity in the loading coil spacing than was necessary in non-repeatered circuits, and a corresponding reduction in the coil inductance deviations. This latter requirement meant that the new coils must have a much greater resistance to the magnetizing effects of superposed steady and transient line-currents, especially since exposure to lightning surges had to be accepted as a normal service experience.

The new requirement for high magnetic stability in the loading coils was met by using short air-gaps at diametrically opposite points in their toroidal-type 65-permeability iron-wire cores. This construction feature also resulted in a substantial reduction in (but not the elimination of) telephone transmission distortion caused by "telegraph flutter" phenomena, when composite telegraph circuits were superposed on the loaded circuits. The new side circuit and phantom circuit loading coils were coded 550 and 549, respectively. They were a little smaller than the coils which they superseded, and had somewhat higher resistances. The resulting attenuation impairments, however, were negligible in the repeatered circuits. The headpiece includes a 550 coil (the largest coil, designated A).

During the decade that followed the beginning of transcontinental telephony, a large enough quantity of Nos. 549 and 550 coils were installed to load approximately 300,000 circuit miles. In the beginning, the improved loading was concentrated on parts of a proposed backbone-network of repeatered 165-mil lines. Soon it became apparent from: (a) the continuing

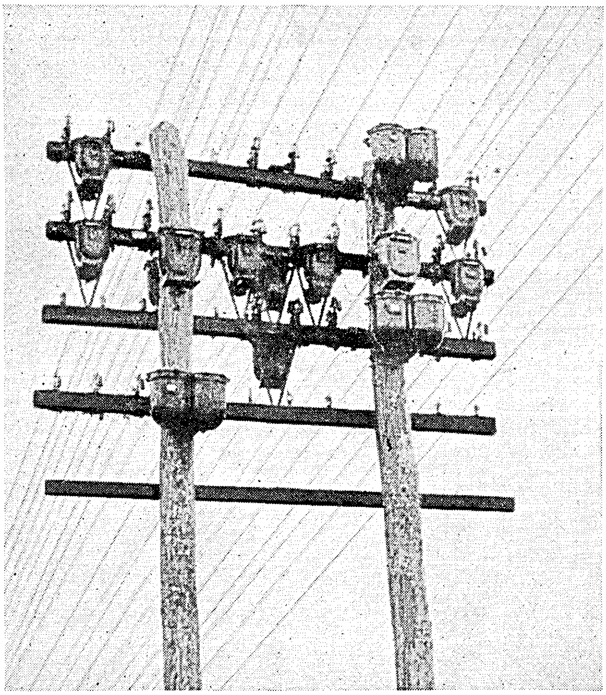


Fig. 4—Typical installation of high stability open-wire loading coils. Note the three phantom loading unit combinations in 3-coil cases supported on the poles; other individually potted side-circuit and phantom coils are supported on the crossarms.

development work on telephone repeaters, repeater circuits and auxiliary apparatus and transmission networks, and from (b) field experiments supplemented by theoretical cost studies, that non-loaded 165-mil lines with additional repeaters would have much more satisfactory transmission characteristics than repeatered loaded lines, and would be less expensive. The principal transmission advantages were: (1) the practicability of securing materially lower net losses, in consequence of the effect of the higher velocity of transmission in reducing echo-current disturbances, (2) more uniform attenuation and impedance characteristics under varying weather conditions, (3) reduced delay-distortion because of the more uniform velocity-frequency characteristics at the upper speech frequencies, (4) complete elimination of telegraph-flutter impairments, and (5) better quality of speech transmission by the effective transmission of a much wider frequency-band. In this latter connection, it is of interest that the transmission band which was effectively transmitted over the loaded, repeatered, transcontinental circuits ranged from about 350 to 1250 cycles, defining the band as that between the lowest and highest frequencies whose transmission was not more than 10 db higher than that of the 1000-cycle transmission. At the higher frequencies, the line losses and the loading coil losses piled up so as to effectively suppress transmission. The excess losses at the low voice-frequencies were due to the line terminal apparatus and the repeater auxiliary apparatus.

The rapidly growing appreciation of these advantages led initially to a curtailment in the installation of new loading on 165-mil circuits, and subsequently to the removal of the existing loading and the installation of additional repeaters.[h] By this time, the vacuum-tube repeater had been accepted in its own right as an independent instrumentality for improving transmission.

On 104-mil lines, the economic competition between loading and repeaters was much closer than that on 165-mil line, and for a period of several years the aggregate mileage of loaded 104-mil lines increased substantially while the mileage of loaded 165-mil lines decreased at an accelerating rate. In this connection the transmission disadvantages of loading on repeatered 104-mil lines were not so serious as those on repeatered 165-mil lines, partly because of the much shorter lengths involved and partly because of their more stable transmission performance under varying weather conditions.

During the early 1920's, the commercial exploitation of open-wire carrier telephone and carrier telegraph systems became an increasingly important factor in the removal of loading from open-wire lines.

About 1924, the practice of installing new loading on 104-mil lines was stopped in order to increase the plant flexibility for the more extensive use of repeaters and of carrier systems, and accordingly the production of new open-wire loading coils was discontinued. The removal of the existing loading, however, was not completed until about 1934.

(5) High Stability Type Coils for Coarse-Gauge Toll Cables

The use of improved telephone repeaters started on a small-scale basis on loaded coarse-gauge circuits along the Boston-Washington route even

[h] The unloading of the original transcontinental circuits was completed early in 1920. The net loss was reduced from about 20 db to 11 db, and the width of the effective transmission band was doubled.

before the loaded, repeatered transcontinental open-wire circuits were ready for service. To permit more satisfactory repeater operation in the remaining loading complements of these cables, and in many more coarse-gauge toll cables which were installed during the next few years, a series of high-stability type of loading coils was standardized during 1916.[i] The coils in this series used air-gaps in their 65-permeability, iron-wire cores. Since the availability of satisfactory types of telephone repeaters had reduced the need for 10-gauge conductors, the new coils were "compromise" designs, suitable for either 10 ga. or 13 ga. conductors. Accordingly they were intermediate in size between the two different series of coarse-gauge cable coils that were developed in the 1911–1913 period.

These new loading coils remained standard for toll cable uses for only a few years. The practice of installing 10 ga. and 13 ga. toll cables substantially stopped before 1920, because theoretical studies of the possibilities of improving repeaters and loading systems were indicating that it should ultimately be possible to use repeatered, loaded 19-gauge or 16-gauge conductors for spanning the longest distances likely to be involved in the long-distance cable plant. The use of 4-wire repeatered facilities became a very important objective in the new development plans, making necessary an intensive development of transmission equalizing and regulating networks and practices.

(6) Compressed Powdered-Iron Core Loading Coils for Repeatered and Non-Repeatered 19 and 16-Gauge Toll Cables

6.1 *Compressed Powdered-Iron Core-Material*

General

This was the first new loading coil core-material to be developed since the establishment of the first loading standards. Many other possibilities had been considered on a number of occasions, notably silicon steel in fine-wire form, but no core-material had been discovered that was superior to the 65-permeability iron-wire in its major performance characteristics.

The compressed powdered-iron development was the pioneering beginning of an entirely new and very important art in the design of high-stability, low-loss, magnetic core-materials. It was started by the Engineering Dept. of the Western Electric Co. as an independent project, alongside the basic research work on various phases of transcontinental telephony, and reached the first stage of commercial fruition during 1916. An important objective was to obtain much better magnetic stability than that of iron-wire cores, and at a lower cost than that of wire cores having series air-gaps. Also, from

[i] Additional information regarding this development is given in References (8) and (9).

the manufacturing standpoint it was desirable to have a better control of the magnetic properties of loading coil core material, and to be free from limitations on quantity such as were occasionally experienced in obtaining iron core wire from outside manufacturers. These particular difficulties were very serious during the First World War in consequence of the greatly increased demand for loading coils and the impossibility of securing an adequate supply of diamond dies for drawing the 4-mil core wire. (In normal times, all dies were imported from Europe.) Incidentally, the supply limitations on diamond dies made it necessary to permit the use of over-size core wire even though this resulted in an impairment in transmission performance due to the abnormal (eddy current) core losses. Fortunately, the compressed powdered-iron core material became commercially available in time to be of great value in helping the Western Electric Co. to increase the output of loading coils.[j]

The success achieved in this development subsequently led to the application of the compressed, insulated, magnetic-powder technology to magnetic alloys, for use in the cores of loading coils and of other types of coils used in various types of transmission networks, including electric wave filters. Initially worked out for voice frequency applications, the new technology expanded to become an important factor in the design economy of carrier and radio transmission systems. The low eddy current losses made possible in large part by the use of very small, insulated, magnetic particles, were inherently important elements in the high frequency applications. Following its development in the United States, the compressed, magnetic powder core, in one form or another, spread to Europe, and became important in world wide communications.

The prior art is of historical interest, in that experiments with finely divided magnetic particles had extended over a period of several decades. As an early example, Oliver Heaviside described in his "Electrical Papers" some work on magnet cores with magnetic powder embedded in wax. It is also of interest that during the Bell System pioneering efforts to obtain satisfactory loading coil cores, considerable experimental work was done (1901) on magnetic oxide cores involving high temperature heat treatments of loosely formed iron wire or iron tape core structures in an oxygen atmosphere. (U. S. Patents Nos. 705,935, and 705,936, July 29, 1902.)

[j] The total output, however, could not be increased sufficiently fast to meet the high 1917–1918 demand for loaded facilities. This resulted in a temporary practice of what came to be known as "omitted-coil loading" on a substantial mileage of toll cable facilities. In the initial installation of loading on these particular facilities, the coils were placed at alternate load points along the line;—for example, at 12,000 ft. spacing instead of the standard 6,000 ft. spacing. The resulting transmission impairments were accepted as being tolerable under war emergency conditions. As soon as practicable, however, the "omitted loads" were "filled in," so that shortly after the end of the war the coil spacing conformed to the established standard practices.

None of the early work by these and other investigators, however, had resulted in commercially usable iron powder cores. It was not until the Western Electric concept of pressures sufficient to deform the magnetic particles, and of an insulating medium of such character as to withstand these high pressures and provide an exceedingly thin insulating film between particles, was developed, that commercially usable results were obtained.

The Western Electric compressed powdered-iron development was carried out in two distinct steps, one after the other, using the same basic magnetic material. In the early work, consideration was given to the use of chemically produced iron powder. Then, mainly for cost reasons, the development

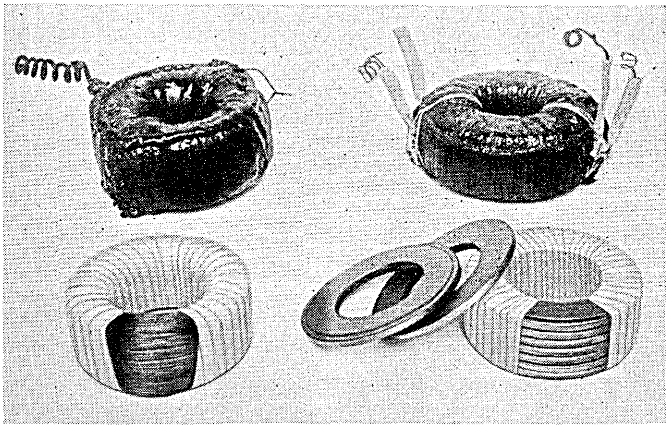


Fig. 5—Early cable loading coils and their cores. At left: Iron wire core coil; core has insulating-binding tape partly removed to show core construction. At right: Compressed iron powder core coil; 7-ring core has binding tape partly removed. Also 2 individual core rings.

efforts were concentrated upon the use of electrolytically deposited iron which was processed so that it could easily be ground into particle sizes of the required fineness. The iron particles were thoroughly mixed with suitable insulating material and then moulded into thin core rings of desired over-all dimensions under moulding pressures of about 100 tons per square inch. These core rings were stacked vertically and bound with insulating tape for use as loading coil cores.

In the first commercially usable product, the iron-powder was annealed prior to the insulating and ring-moulding operations. When a higher-grade loading coil core material became necessary for reasons subsequently discussed, the iron-powder annealing process was omitted and a larger amount of insulation was mixed with the magnetic material, which changes resulted in a substantial reduction of effective permeability. Other process differences were also involved. These two magnetically different, compressed loading coil core-materials were sometimes known as "soft-iron dust" and "hard-iron dust" or as "compressed annealed iron-powder" and "compressed unannealed iron-powder," respectively. In the Speed-Elmen classic paper[13] on "Magnetic Properties of Compressed Powdered Iron," these two materials are referred to as "Grade A" and "Grade B", respectively. A still lower permeability core material, known as "Grade C", was developed primarily for use in carrier frequency inductance coils. In this material, a larger amount of particle insulation than that in the "A" and "B" grades was used, and the average size of particle was smaller.

In commercial production the "Grade B" cores consisted of a mixture of 90% unannealed powder and 10% annealed powder, the latter component being included to obtain the desired value of permeability, and to increase the mechanical strength of the core rings.

The different magnetic characteristics of the "A" and "B" grades of compressed, powdered iron were basic factors in the evolution of the loading practices for the new loading coils that used them in their cores. A brief review of these practice differences follows, and includes some additional general data regarding the coils themselves.[(k)]

6.2 *Compressed, Annealed, Powdered-Iron Core Loading Coils*

Referring to Table III, it will be noted that cores using this new material had nearly the same effective volume-permeability as the cores of 95-permeability iron-wire. By using similar-size cores, and closely similar windings, it was found possible to obtain effective resistance-frequency characteristics close to those of the standard small-gauge cable loading coils using 95-permeability wire cores, as typified in the 508 coils of Table I, and corresponding grades of side circuit and phantom loading coils. Also the new potting developments were minimized.

An outstanding service advantage of the new "soft-iron dust" core loading coils was in their very high stability of residual inductance, by virtue of the self-demagnetizing action of the very large number of very small series air-gaps in the cores. After a temporary exposure to magnetization by abnormally large superimposed currents that might be caused by accidental grounds on superposed d-c signaling circuits or by induction from outside sources (lightning, power-line shorts or grounds), the coil inductance would return to within a few per cent of the initial value. On the other hand, after extreme exposure to strong magnetic shocks the residual inductance in the 95-permeability wire-core coils might be as much as 40% below the initial inductance, the high retentivity of the magnetic circuit being an important factor in this performance.

(k) Some additional detailed data regarding these two series of loading coils were published in Reference (8).

The new coils also were considerably better than the 95-permeability wire-core coils in the following important features:

(a) Their susceptibility to changes in inductance and effective resistance during service intervals involving the superposition of steady d-c signaling currents;

(b) Their susceptibility to the transient magnetizing effects of superposed composite-telegraph currents, i.e., "telegraph flutter".

The relative performance characteristics, above described, resulted in the "soft-iron dust" core coils superseding the standard 95-permeability wire-core coils in the fields of use in which these older standard coils had been used. As an important example, the original standard 508 coil, used principally for medium loading in exchange cables, was superseded in 1916 by the 574 coil, which remained standard for about a decade.

The telegraph-flutter characteristics, Item (b) above, prevented the new coils from being used generally in place of 65-permeability wire-core coils on toll cables quads having all four wires composited for grounded telegraph operation. However, for a few years there was a "compromise" practice of combining "soft-iron dust" side circuit loading coils with 65-permeability wire-core phantom loading coils in 19 and 16-gauge toll cable projects where the needs for superposed grounded-telegraph operation could be satisfied by compositing the phantoms, and the demands for repeatered facilities could be met by limiting repeater operation to the side circuits. In this special loading setup, the transmission distortion by "telegraph flutter" was controlled in the phantoms, and was completely avoided in the side circuits because the grounded telegraph currents, flowing in parallel through the side circuit coil windings, neutralized each other's effect in magnetizing the cores. With respect to regularity in circuit impedance-frequency characteristics in relation to repeater gains, the high residual-inductance stability of the soft-iron dust-core loading coils made them distinctly preferable to the 65-permeability wire-core coils in the repeatered side circuits.

During 1917–1918, when the subsequently described work on improved loading systems for long repeatered toll cables got well under way, theoretical studies of the use of soft-iron dust core loading coils on such facilities disclosed seriously objectionable non-linear transmission distortion that had not been bothersome on short circuits. This was due to the relatively large hysteresis losses in the loading coil cores, which cause the effective resistances of the coils and the circuit attenuation loss to increase appreciably in magnitude as a function of line current amplitude. The effects of these losses are much more serious in the repeatered circuits, because of the larger line currents, and because of the much greater circuit lengths. Since the hysteresis losses also vary in direct proportion to the telephone frequency, the resultant coil-resistance increments and attenuation increments are greater at the high-speech-frequencies than at low frequencies.

As none of the other energy losses in the core or winding vary with line current strength, it became convenient to consider the attenuation loss caused by hysteresis as an "excess" loss, when referred to the attenuation that would result if the hysteresis should be vanishingly small or zero. The theoretical studies above mentioned not only showed the "excess" attenuation in long loaded, repeatered, circuits to vary as a function of the magnitude of the input current and frequency, as above noted, but also showed it to vary as a function of the length of repeater section, the position of the repeaters in the line, the weight of loading, and the over-all circuit length. Since it is not possible to offset the effects of loading coil hysteresis by means of distortion corrective or equalizing networks at repeater stations or circuit terminals, the piling up of the "excess" losses along the line in very long loaded, repeatered, circuits could reach values that would be large relative to the desired over-all working equivalent, if the individual loading coils should have large hysteresis losses, as did the soft-iron dust-core loading coils.

Comparative theoretical studies of the excess loss due to hysteresis effects in 65-permeability core loading coils showed these coils to be greatly superior to the soft-iron dust-core coils in this feature. On the other hand, the wire-core coils were relatively unsatisfactory from the inductance-stability standpoint for use on long repeatered circuits.

6.3 *Compressed, Unannealed, Powdered-Iron Core Loading Coils*

It was very fortunate that the development work on the compressed, unannealed, powdered-iron core-material, previously mentioned, approached commercial fruition at about the time the unsuitability of the soft-iron dust-core loading coils for very long repeatered circuits became apparent.

As noted in Table III, the effective volume permeability of this improved core-material was closely that of the 65-permeability wire-cores. The new standard loading coils using this improved material had cores generally similar in dimensions to those of the older coils used on 19 and 16 ga. toll cables, and their over-all dimensions were sufficiently similar to avoid the need for developing new loading coil cases.

In general terms, the new coils combined the best qualities of the soft-iron dust-core loading coils with the best performance characteristics of the 65-permeability wire-core loading coils. Actually, they were much better than the soft iron-dust core coils with respect to stability of residual inductance, and susceptibility to magnetization by superposed steady currents. In these respects they were also substantially superior to the 65-permeability wire-core loading coils. However, they were not quite so good as the low permeability wire-core coils with respect to hysteresis losses and telegraph flutter transmission impairments. On the other hand, they were substantially as good in their effective resistance-frequency characteristics. The above summarized advantageous electrical and magnetic properties resulted during 1918 in the standardization of the new compressed, annealed, powdered-iron core loading coils for general toll cable use in place of the 65-permeability wire-core coils and soft-iron dust-core coils. The availability of the improved loading coils quickly stopped the previously mentioned temporary, compromise practice of using soft-iron dust side circuit loading coils in combination with 65-permeability wire-core phantom loading coils.

After the hard-iron dust-core loading coils became commercially available, the remaining development work on improved toll cable loading systems described in the following pages was in terms of these coils.

(7) New Loading Systems for Repeatered 19 and 16 Ga. Toll Cables

7.1 *General*

The basic problems of learning how to use telephone repeaters most advantageously on long cable circuits began to receive serious attention soon after the completion of the open-wire transcontinental telephony project, along with the repeater development work that ultimately resulted in the obsolescence of open-wire loading, as previously mentioned. During the decade or more of intensive, continuous development activity on the repeatered toll cable problem that followed, it was found highly advantageous to work loading and repeaters together as equal partners in a team, each making its contribution according to its own nature. The important contributions of loading were the substantial reductions of attenuation and of frequency distortion at a cost (for voice-frequency transmission) much lower than the cost of the additional repeaters and the distortion corrective-networks which would have been required on non-loaded cables. Incidentally, the use of loading substantially simplified the solution of the important equalization and regulation problems.

The attainment of the good working partnership between loading and repeaters involved the development of new loading systems having substantially improved transmission characteristics, and the development of improved repeaters and improvements in repeatered circuits, including equalizing networks and arrangements for controlling the cable transmission-performance changes that result from seasonal and daily changes in temperature. A classic report on these related developments is given in a paper[14] by A. B. Clark, entitled, "Telephone Transmission over Long Cable Circuits."[(1)] The present discussion is primarily concerned with the features of the improved loading that were essential to satisfactory transmission-performance in long repeatered cable circuits.

[(1)] Reference (15) is also of interest with respect to engineering aspects.

The compressed, unannealed, powdered iron-core loading coils previously described were found to be satisfactory with respect to inductance stability and other properties, including hysteresis effects for use in the improved loading systems. New smaller inductance values were necessary, however, for the coils used in the longest circuits. It was of course necessary to control the geographical spacing deviations, and the factory deviations in cable capacitance and in the loading coil inductances, so as to obtain a satisfactory degree of "regularity" in the impedance characteristics of the loaded circuits.

7.2 *Transmission Limitations of Existing Loading Systems*

(a) *General:* As the lengths of loaded cables progressively increased beyond those involved in the establishment of the first standard loading systems, certain transmission effects which were initially unnoticed or comparatively unimportant became very noticeable as objectionable transmission impairments. By increasing the electrical lengths of the circuits, the use of repeaters greatly aggravated these impairments and it became very desirable to correct them so far as feasible at their source. Complex problems thus arose in providing better quality of transmission over much greater distances.

The impairments referred to above were directly related to the band width of frequency transmitted by the line, which is determined by the cut-off frequency, and to the velocity of transmission. They are discussed briefly in the following paragraphs.

(b) *Attenuation-Frequency Distortion:* At frequencies above about 70% of the theoretical cut-off frequency, the attenuation increases with rising frequency at a continuously accelerating rate, in consequence of the accumulation of the effects of internal reflections at the individual loading points.

At lower frequencies where these so-called "lumpiness of loading" effects are not of dominating consequence, the attenuation increases with rising frequency are largely due to the energy losses in the loading coils. (Usually the eddy-current losses in the cores are the most important component loss, since they are proportional to the square of the frequency.) It thus happens that the attenuation losses may pile up in long loaded cables in such a way as to substantially suppress the transmission of the higher-frequency components of speech, even when the attenuation losses are tolerable at lower frequencies. In consequence, the width of the transmission band which is effectively transmitted over a long loaded cable becomes narrower and the quality of transmission progressively deteriorates as the circuit length increases, unless suitable auxiliary equalizing networks and additional repeaters are utilized. The large amount of attenuation-frequency distortion which occurred in the first transcontinental telephone circuits was previously commented upon.

The attenuation-frequency distortion in long loaded cables can be reduced by raising the theoretical cut-off frequency. As a secondary factor, the reduction of core losses in the loading coils is advantageous.

(c) *Velocity Distortion:* This became noticeable as peculiarly disturbing, transient distortion in the intervals when spoken syllables were building up or dying down, prior to, or after, their steady-state transmission. It is most disturbing at the upper speech-frequencies where the steady-stage velocity of wave propagation varies at an accelerating rate as the cut-off frequency is approached. It can be particularly disturbing in very long

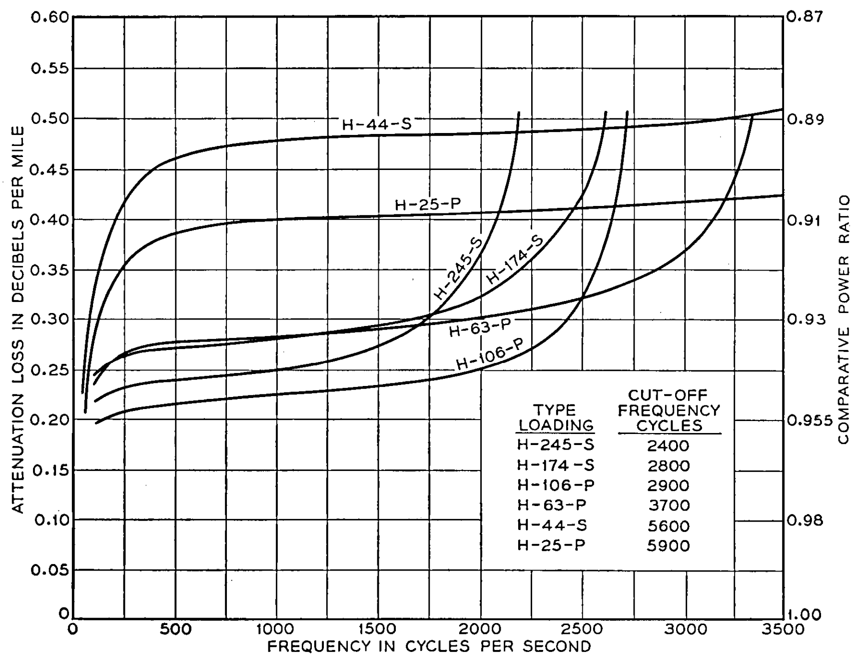


Fig. 6—Attenuation frequency characteristics toll cable loading.

circuits where repeaters are used to reduce the over-all loss, and in extreme cases it could make the circuits unusable for commercial service. These transient distortion impairments may be reduced by raising the loading cut-off frequency, and by increasing the velocity of wave propagation. The loading design changes that were made in these features gave a satisfactory control of the velocity distortion in the longest loaded cables used commercially, without requiring the use of velocity-distortion corrective networks in the lines or at repeater stations.

(d) *Echoes:* Echo effects also were found to be potentially limiting factors in providing satisfactory transmission-performance over long loaded cable circuits. They are due to the transmission of reflected energy from points of impedance irregularity. They are troublesome factors whenever the time of transmission between the point of reflection and the disturbed subscriber is appreciable, especially when the use of repeaters prevents the attenuation of the reflected energy from being negligible in magnitude. Increasing the velocity of wave propagation is a sure procedure for reducing echo effects. Control of the impedance irregularities which cause them is also important, but cannot be carried to a sufficiently fine degree in long, low-velocity circuits. For the satisfactory control of echo effects in very long, four-wire type, high-velocity, (H44-25) loaded circuits, the use of auxiliary devices known as "echo suppressors" was found to be very advantageous.[16]

(e) *Basic Networks and Filters:* The previously mentioned basic networks[12] which simulate the impedances of the loaded circuits are vitally necessary features of the balancing lines that are used with the repeaters employed on two-way speech circuits. Relatively simple types of networks provide satisfactory impedance simulation up to a high fraction of the cut-off frequency, in loaded circuits having regular coil spacing and uniform coil inductances. At frequencies near the loading cut-off, however, it is not feasible to provide basic networks which satisfactorily simulate the impedances of the loaded circuits. Consequently, in order to avoid large repeater unbalances which would cause objectionable singing at these frequencies, it is necessary to associate with the repeaters low-pass type electric wave filters[17] which have cut-off frequencies appreciably lower than the loading cut-off frequencies. Usually this cut-off frequency differential is of the order of 10% or more of the loading cut-off. The reduction of transmitted band width caused by these filters aggravates the frequency attenuation distortion effects previously discussed. This use of filters with the repeaters thus became a contributory factor in the need for raising the loading cut-off frequency in long repeatered circuits.

By reducing the transmitted frequency band width, however, the filters used with the repeaters have favorable effects in reducing the high-frequency velocity distortion caused by the loading. The filters used with the repeaters on long H44-25 circuits, subsequently described, had cut-off frequencies substantially below the loading cut-off frequencies primarily for the purpose of controlling velocity distortion impairments.

7.3 *Improved Loading Systems*

(a) *General:* To sum up the foregoing, the theoretical analyses of the limitations of the standard toll cable loading pointed definitely towards an increase in cut-off frequency and in the velocity of wave propagation. Since the state of the art was such that theoretical studies[14, 18] alone could not determine the magnitudes of the changes that would be required, extensive investigations of experimental installations also became necessary. For cost reasons, it seemed desirable that the improved loading should be used at standard "heavy loading" spacing, i.e., 6000 ft. In consequence, the increase in cut-off was proportional to the increase in velocity. Also, there was a proportional reduction in the nominal impedance, accompanied by an increase in the unit-length attenuation.

TABLE IV
LOADING SYSTEMS—SMALL GAUGE REPEATERED TOLL CABLES

Item No.	Loading [a] System	Circuit	Coil [b] Code No.	Nominal Impedance (ohms)	Theoretical Cut-off Frequency (cycles)	Nominal Velocity (mi/sec.)	Attenuation Loss[c] (db/mile at 1000 Cycles)		Maximum[d] Geographical Length (miles)
							19 A.W.G.	16 A.W.G.	
(1)	H174-106	Side	584	1550	2800	10000	0.28	0.16	500
(2)	"	Phantom	583	950	2900	10000	0.22	0.13	500
(3)	H44-25	Side	590	800	5600	19000	0.48	0.25	2000
(4)	"	Phantom	589	450	5900	20000	0.40	0.21	2000
(5)	H174-63	Side	584	1550	2800	10000	0.28	0.16	500
(6)	"	Phantom	587	750	3700	13000	0.28	0.16	1500
(7)	H245-155	Side	582	1850	2400	8000	0.25	0.16	250
(8)	"	Phantom	581	1150	2400	8000	0.20	0.12	250

Notes: (a) Nominal coil spacing is 6000 feet in cable having a capacitance of 0.062 mf/mile in the side circuits and 0.100 mf/mile in the phantom circuits.
(b) The code numbers of the first standard compressed, unannealed, powdered-iron core coils used in the loading systems.
(c) These attenuation values apply at 55°F. Under extreme high or low temperature conditions, the actual attenuation may be approximately 12 per cent larger or smaller, due principally to changes in conductor resistance with temperature. In long repeatered cable circuits these variations of attenuation with temperature require special corrective treatment by means of automatic transmission regulators.
(d) These particular length-limitations were set by velocity-distortion effects. By using velocity-distortion corrective networks in the lines or at repeater stations, it would have been possible to extend the circuit lengths beyond the listed limits, provided also that adequate steps could be taken to control echo currents. Under actual service conditions, however, echo currents might limit the circuit lengths to considerably lower values than those listed above, depending upon the grade of balance of the lines and the permissible overall loss.

The transmission development work resulted in the standardization of two new phantom-group loading systems, designated H174-106 and H44-25[m] in Table IV. Several years later (1923), a new H63 phantom

[m] The letter-number loading designations used in Table IV, and in the remainder of the text, were simplifications adopted for general use in 1923. The letter prefix symbolizes the geographical spacing in feet; the numbers correspond to the nominal inductances (in millihenrys) of the associated side circuit and phantom loading coils, in the sequence noted. The letter "H" designates "Heavy" loading spacing. In the early days this was about 1.25 miles; later it became 6000 ft.

loading was substituted for H106 phantom loading, to provide H174-63 phantom-group loading as a successor standard for H174-106 loading.

(b) *H174-106 Loading:* The development work on H174-106 loading preceded that on the H44-25 system. By using available standard "medium" loading coils at standard "heavy" spacing, the transmission velocity and the cut-off frequency were raised to values about 20% higher than those of the H245-155 loading which had been, by far, the most widely used loading on 19 and 16 gauge toll cables. The new combination of inductances and spacings became widely known in the early installations as "medium-heavy, high cut-off" loading. This designation called attention to the first change in standard loading cut-off frequency since the establishment of the initial loading standards in 1904.

When used in conjunction with improved repeaters, the H174-106 loading enabled satisfactory transmission to be obtained over circuits about twice the length of the longest H245-155 repeatered circuits which were satisfactory from the transmission standpoint. In the beginning, H174-106 loading was extensively used on 4-wire repeatered circuits. After the new transmission systems using H44-25 loading came into general use, H174-106 (and H174-63) loading was largely restricted to short haul two-wire circuits.

The 1917 trial installation tests showed H174-106 loading to be a substantial step forward in the struggle to extend the transmission range in repeatered 19 and 16-gauge toll cables, but far from a big enough step to satisfy the transmission requirements in very long cables such as the New York-Chicago cable project which had been accepted as a definite development objective. The continuing studies which considered other combinations of lower inductances and of standard and new spacings ended in the decision to standardize H44-25 toll cable loading.

(c) *H44-25 Loading:* Although these inductance values had been used for impedance matching loading on entrance cables, they had not previously been used on toll cables. The initial designation for the improved loading was "extra-light, very high cut-off " loading. The cut-off frequency and transmission velocity were about twice as high as those in H174-106 loading, and the nominal impedance was 50% lower.

H44-25 was necessary for the longest repeatered circuits. It was developed primarily for use on 19-gauge 4-wire circuits in which large repeater gains could be obtained by the repeaters in the one-way paths, to offset the relatively high bare-line attenuation. The first installation of H44-25 loading was made during 1919 on circuits in the New York-Philadelphia-Reading Cable. Trial service of a complete four-wire system, including new regulating and equalizing arrangements, and an echo suppressor, started during 1923. In October 1925, commercial service between New York and Chicago started over a H-44-25 four-wire repeatered system.

It had taken a long time to work out the necessary improvements in the repeaters and their associated distortion-corrective networks, and to learn how to use regulating repeaters to control satisfactorily the very large transmission changes that resulted from temperature variations in long circuits.

During the mid 1920's aerial cable came into use extensively along new cross-country routes where underground cable conduits would have been unduly expensive. Such cables, however, presented added difficulties and expense in transmission regulation since the temperature range variations are about three times as great as in underground cable. During 1930 the use of buried cables started. With respect to transmission regulation and service continuity, they compare well with cables in underground conduit and are less expensive, when the number of cables along the route is small. However, buried toll cable generally tends to be more expensive than aerial cable. At the end of 1949 the aggregate wire mileage in buried toll cables was nearly one-third of that in aerial toll cables.

A substantial amount of H44-25 loading was also used on 16-gauge 2-wire repeatered facilities, for circuit lengths intermediate between the transmission limits of H174-106 and H174-63 loading and lengths where echo-impairment difficulties made necessary the use of 19-gauge 4-wire facilities. In such two-wire circuits, very good line-balance was of course required at the intermediate repeaters. An important economic factor in this "medium" long-haul practice was the 50% lower loading cost per unit of facility length of the two-wire circuits. This intermediate-length usage of 16-gauge 2-wire circuits, however, tapered off in the long-distance plant of the A.T.&T.Co. during the late 1920's, so as to obtain the important plant flexibility and operating advantages that were inherent in the general use of 19-gauge 4-wire circuits for medium-haul and long-haul toll cable facilities.

Notwithstanding the 2:1 ratio in loading cut-off frequencies, the width of the frequency-band transmitted over long H44-25 circuits was not much wider than that transmitted over the much shorter H174-106 facilities. This was largely due to the filters which were used to suppress the upper half of the H44-25 transmission band, primarily for the purpose of reducing the very serious velocity-distortion transmission impairments that would have otherwise resulted in very long circuits. This suppression of the higher frequencies also eased the transmission equalization and regulation problems.

The attenuation-frequency distortion characteristics of the old and new toll cable loading are shown in Figs. 6 and 7.

Figure 6 (p. 175) shows the attenuation in db per mile. Figure 7 (p. 180) gives the bareline attenuation-frequency curves under specified circuit length and repeater conditions in which the total 1000-cycle attenuation is 10 db. The effect on frequency distortion of raising the loading cut-off frequency, and of using distortion corrective-networks, is clearly indicated.

(d) *H174-63 Loading:* Before concluding this summary of the basic development work on improved loading for long repeatered circuits, it is opportune to include some detailed information on the previously mentioned improvement of phantom loading for the so-called "medium-heavy" loading system, i.e., the combination of new H63 phantom loading with the existing H174 side circuit loading to constitute the H174-63 phantom-group loading system, data for which are included in Table IV. The practical importance of this development was in the substantial reduction it permitted in loading costs. It marked the beginning of commercial use (1923) of phantom loading coils having cores similar in size to those of the associated side circuit coils. The cost savings resulted directly from the coil size-reduction, and from the increased size of loading complements that could be placed in standard-size

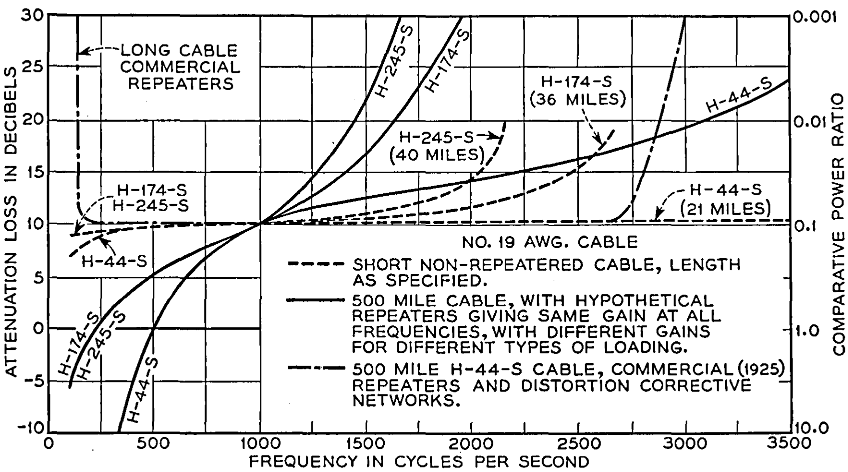


Fig. 7—Attenuation frequency characteristics of short and long loaded toll cable circuits having a net attenuation loss of 10 db at 1000 cycles per second.

pots, and from the simplification of pot assembly and cabling arrangements. The choice of phantom coil inductance (63 mh) was such as to make the phantom circuit 1000 cycle attenuation the same as that on the associated side circuits, without increasing the side circuit attenuation above that in H174-106 loading. This reduction in the loading inductance substantially improved the phantom circuit's transmission performance-characteristics by virtue of the substantial increase in cut-off frequency and transmission velocity, and made the repeatered phantom circuits electrically suitable for much greater distances than their side circuits. This superiority was seldom utilized, however, because of the practical operating flexibility advantages inherent in the established practices of using the associated side circuits and phantoms interchangeably between the same operating or switching centers.

In due course, the sizes of the other toll cable phantom loading coils were reduced to side circuit coil size, for cost-reduction reasons. The resultant economies were large relative to the value of the small attenuation impairments which resulted from this change.

(8) New Telegraph Systems for Loaded Cables

At several points in this review references have been made to the transient telephone transmission impairments which resulted from the operation of separate, superposed, grounded telegraph circuits over the individual line-wires of the loaded telephone circuits.[n] These "composite" telegraph systems had originally been developed for use on non-loaded open-wire lines

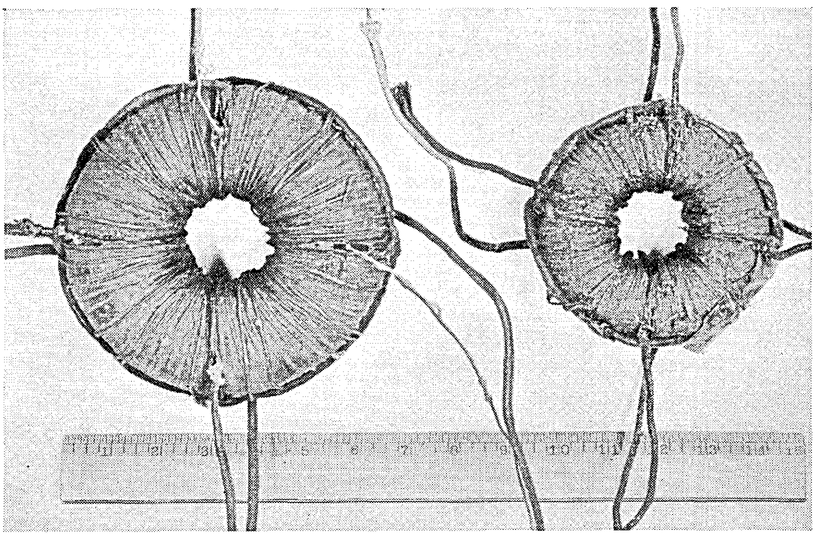


Fig. 8—Reduction of cable phantom coil size to that of associated side circuit coils. Large No. 581 Phantom Coil (106 m.h.) at left. Small No. 587 Phantom Coil (63 m.h.) at right. N.B. These phantom coils had compressed unannealed iron-powder cores.

and consequently required the utilization of telegraph currents of very large amplitude relative to the telephone currents.

The great extensions in the lengths of 19-gauge repeatered loaded cable that were to be expected from the use of the improved loading and repeaters, previously considered, put great emphasis upon a much better control of the telegraph-flutter interference, and led to the development of improved cable telegraph systems, along with the improved telephone transmission systems.

In one of these, known as the "Metallic Polar Duplex Telegraph System,"[19] the superposed telegraph current was of the same general order

[n] Reference (10) gives a comprehensive discussion of telegraph-flutter phenomena.

of magnitude as the telephone current. This permitted a much greater reduction of the telegraph-flutter impairments than that which could have been obtained at a reasonable cost by using larger coils. A disadvantage of this system, however, was that it halved the number of superposed telegraph circuits per toll cable quad.

The other new telegraph system was a voice-frequency carrier system[20] which became commercially available during 1923 soon after the direct current metallic system just mentioned. This provided a total of 10 independent telegraph channels over the side circuits of special groups of H174-63 4-wire circuits used exclusively for telegraph service, and consequently there could be no telegraph-flutter reactions on telephone transmission. On the other hand, the need for controlling intermodulation effects among the associated carrier-telegraph channels imposed limits upon the allowable non-linear distortion in the loading coils. The loading coils then standard (having compressed, unannealed, iron-powder cores) satisfactorily met the requirements, and with a greater margin than did the repeaters which were standard at that time. Interference considerations, however, prevented the general use of the metallic polar telegraph systems on loaded cable pairs used for carrier telegraphy.

The voice-frequency carrier-telegraph system soon became the most widely used telegraph system in the long distance toll cables. It did not require special facilities, but made effective use of the whole frequency-band provided for voice telephony. The initial number of channels was expanded to 12 on H174-63 facilities, and subsequently to a total of 24 channels by using the wider-band H44-25 facilities, previously described. The present-day system is limited to 18 channels, however, in order to permit the ready interconnection of loaded cable and broad-band telephone facilities in tandem. A detailed account of these and other telegraph improvements is given in a 1940 B.S.T.J. paper[21].

The strong-current, composite-grounded d-c telegraph system is very seldom used on modern toll cables. There is, however, a considerable use of the strong-current grounded d-c telegraph system on a simplex-phantom basis. Under this service condition, the telegraph current does not magnetize the loading coil cores and consequently there is no telegraph-flutter interference with telephone transmission. The metallic polar d-c telegraph system is usually limited to a few voice repeater-sections, because of the modern severe limits on telegraph-flutter impairments upon telephone transmission, and partly for economic reasons.

The improvements in telegraphy over loaded toll cables also included arrangements which were developed in 1922 for the purpose of reducing interference between d.c. telegraph circuits when superposed by the compositing method on wires of the same loaded cable quad.[22] This interference, known as telegraph crossfire, is mainly due to capacitance coupling between the cable conductors and in the central office equipment. In long circuits, the inductive coupling between the two line windings of each side circuit coil, and among the four line windings of each phantom loading coil, are important factors in the over-all crossfire between the telegraph circuits that are associated with the same cable quad.

(9) Compressed Permalloy-Powder Core Loading Coils

Since the transmission characteristics of the compressed, unannealed, powdered-iron core coils, which became available for general use in toll cables during 1918, were as satisfactory as was expected for the rapidly expanding repeatered toll-cable plant, greater emphasis was placed on cost reduction than on transmission improvement, in the continuing studies of new loading coil design-possibilities.

9.1 *Core-Material Development*

It was inevitable that permalloy[23] should be considered. This remarkable new nickel-iron alloy, invented by Mr. G. E. Elman of Western Electric research department, had important early applications in thin-tape form for continuous loading of deep-sea telegraph cables.

Some early studies and experiments indicated interesting possibilities of using it in thin sheets in non-toroidal type loading coil cores, but the prospects were much more intriguing if permalloy could be made available in compressed-powder toroidal cores. However, the initial experimental results with powdered-permalloy were disappointing. When the processes used in making compressed powdered-iron cores were employed, the permeability was unsatisfactorily low in consequence of the magnetic changes caused by the severe mechanical treatment involved in the embrittlement processes. The development moved forward rapidly after experiments with a physically sturdy type of ceramic-powder insulation for the permalloy particles proved that an annealing treatment after the core rings were pressed could erase the objectionable magnetic effects of the powderizing process and raise the permeability to desirable, high values. A complete account of this very important development is given in an A.I.E.E. paper[24] by W. J. Shackelton and I. G. Barber, "Compressed Powdered Permalloy, Manufacture and Magnetic Properties."

In the form developed for voice-frequency loading coil cores, the effective volume-permeability of the improved core-material was 75, more than twice that of the standard compressed, unannealed, powdered-iron core-material, and the intrinsic permalloy characteristic of very low hysteresis was retained. The combination of magnetic and electrical properties was such that large size-reductions could be made in the loading coils without degrading the important transmission-performance characteristics, relative to those of the standard powdered-iron core loading coils.

9.2 *Permalloy Core Loading Coils*

The coils standardized for toll cable loading were in volume and weight about one-third as large as the superseded types previously described. Coils C and B in the headpiece exemplify the coils under comparison. The direct-current resistances of the new coils were slightly lower than those of the superseded designs, their hysteresis losses were substantially lower, but their eddy-current losses were somewhat greater, because of the higher permeability. In consequence, their effective resistances were more favorable at the important low and middle speech-frequencies, thàn those of the compressed annealed iron powder core coils but not quite as good at the top frequencies.

The stability of residual inductance after magnetization by strong superposed currents was appreciably better than that of the superseded designs. The hysteresis advantages included a substantial reduction in non-linear distortion effects and about a 50% reduction in the transient distortion effects that are unavoidable in the operation of grounded telegraph over composited circuits. Their telegraph-flutter rating was considerably better than that of any of the prior standard cable loading coils, excepting only the large-size "high-stability" type of coarse-gauge cable coils, previously described. (Subdivision 5.)

With the standardization of the permalloy-core loading coils there started the practice of coding the combination of two side circuit coils and the associated phantom coil as a phantom-group "loading unit." The letter "P" was used in the code designations; a code number associated with the code letter recognized the different inductances of the different loading units, the complete codes being P1, P2, etc.

The coil size-reduction resulted in a 2 to 1 reduction in the potting-space requirements and permitted twice as many coils to be potted in standard-size cases, thus reducing potting costs. The cost savings in potted coils ranged from 30 to 40%. Additional savings resulted in the installation costs, including the space costs in the loading vaults in underground cable projects.

The development was very timely, in that the much cheaper coils became available during 1928 for use in the unprecedented expansion of H44-25 four-wire repeatered facilities that started that year and built up to a very high peak during 1930. In the period 1928–1931, over 4,000,000 permalloy-core toll cable loading coils were manufactured for Bell System uses. The lower costs of these coils encouraged the provision of larger circuit-groups in the long-distance cable plant which made possible substantial improvements in the speed of service. This, in combination with excellent transmission performance, greatly increased the demand for service, up to the beginning of the business depression in the early 1930's.

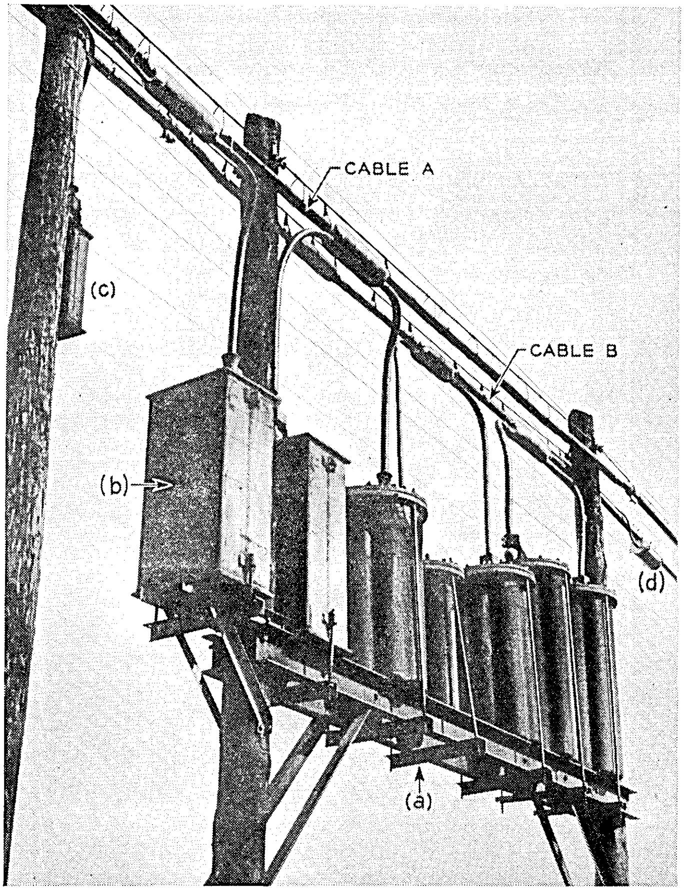


Fig. 9—Installed aerial toll cable loading. Four different methods of supporting the loading-coil cases are used in this installation, which provides loading for two cables: (a) On the platform of a 2-pole "H" fixture. (1 large welded steel case, and 5 large cast iron cases.) (b) A pole balcony supporting a large welded steel case. In this installation, it provides an extension of the "H" fixture. (c) A small welded steel case equipped with brackets, and fastened directly to pole. (d) Clamping a small lead-sleeve case to the main cable and its supporting strand. (This case contains program circuit loading coils.)

Around 1930, improved assembly-arrangements were worked out for the permalloy-core loading units. These involved the assembly of the individual loading units in individual unit-containers, and the code designations were changed. These used the letters "PB", and the same numbers as in the P-type units; the complete designations being P1B, P2B, etc.

The compressed powdered-permalloy core toll cable loading coils remained standard until 1938, when much cheaper and slightly better compressed, powdered molybdenum-permalloy core coils became available, as discussed in subdivision 11.

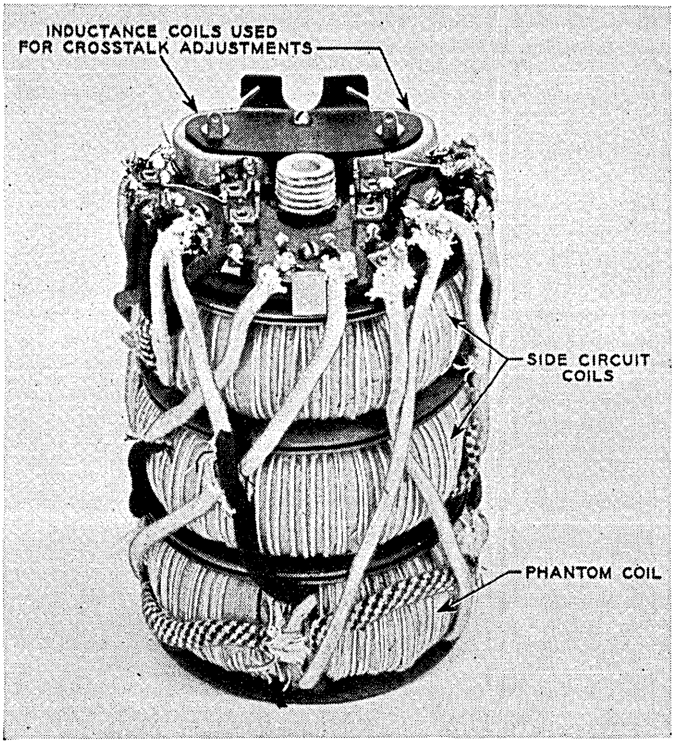


Fig. 10—P-B type loading unit potting assembly prior to placement in unit shielding container. Phantom coil is below the two side circuit coils. Midget inductance coils are used in phantom-to-side crosstalk adjustments.

(10) Improved Loading Standards for Two-Way Repeatered Toll Cables

During the late 1920's considerable attention was given to the improvement of transmission standards made possible by the use of the improved repeaters and the improved loading. This included reductions in the allowable net losses between terminals and improvements in crosstalk performance. Also, steps were taken to obtain better control of attenuation-frequency distortion, the important objective being to provide fairly uniform transmission in the frequency-band between 250 and 2750 cycles per second.[25]

The transmission requirements over the frequency-band just mentioned could be met without undue difficulty on the H44-25 four-wire and two-wire facilities, but were not feasible on the H174-63 facilities then being used mainly on a two-wire basis.

The work on the improved standards problem resulted in the development of the B88-50 and H88-50 toll cable loading systems, data for which are given in Table V.

In effect, this development established a new minimum cut-off standard of about 4,000 cycles per second for loading used on repeatered circuits. Also, as noted in the table, the transmission velocity was increased in H88-50 loading. B88-50 loading, using a coil spacing of 3,000 ft. (i.e., one-half of H-spacing), was originally intended for use in "long" repeater sections. The cheaper H88-50 loading was used on "short" repeater sections, and in consequence some facilities had tandem combinations of the two new types of loading. In the early applications H88-50 loading was used in repeater sections ranging up to about 45 miles in length and the more expensive B88-50

TABLE V
H88-50 AND B88-50 LOADING

Loading System	Circuit	Nominal Impedance (ohms)	Theoretical Cut-off Frequency (cycles)	Nominal Velocity (mi/sec.)	Attenuation Loss (db/mile at 1000 Cycles)	
					19 A.W.G.	16 A.W.G.
H88-50	Side	1100	4000	14500	0.35	0.19
	Phantom	700	4200	15000	0.30	0.16
B88-50	Side	1550	5600	10000	0.25	0.16
	Phantom	950	5900	10500	0.23	0.14

loading was used on longer sections. Later, improvements in the control of crosstalk and special procedures for reducing loading section capacitance-deviations made H88-50 loading suitable for longer repeater sections. During recent years, the voice frequency repeater points have usually been laid out to permit the use of H88-50 loading, and at present there is little use for new B88-50 loading.

New loading coils were developed for the new loading systems, and were coded in the "PB" loading unit series, previously mentioned. This apparatus development included substantial improvements in crosstalk performance obtained by new assembly-methods, and refinements in the crosstalk adjustments and test circuits. These new assembly-methods were applied to all standard toll cable and toll entrance cable loading units and provided flexibility for potting different types of loading units in the same case, and for identification of their terminal leads in the stub cables.

The new H88-50 and B88-50 loading became available for general use during 1932.

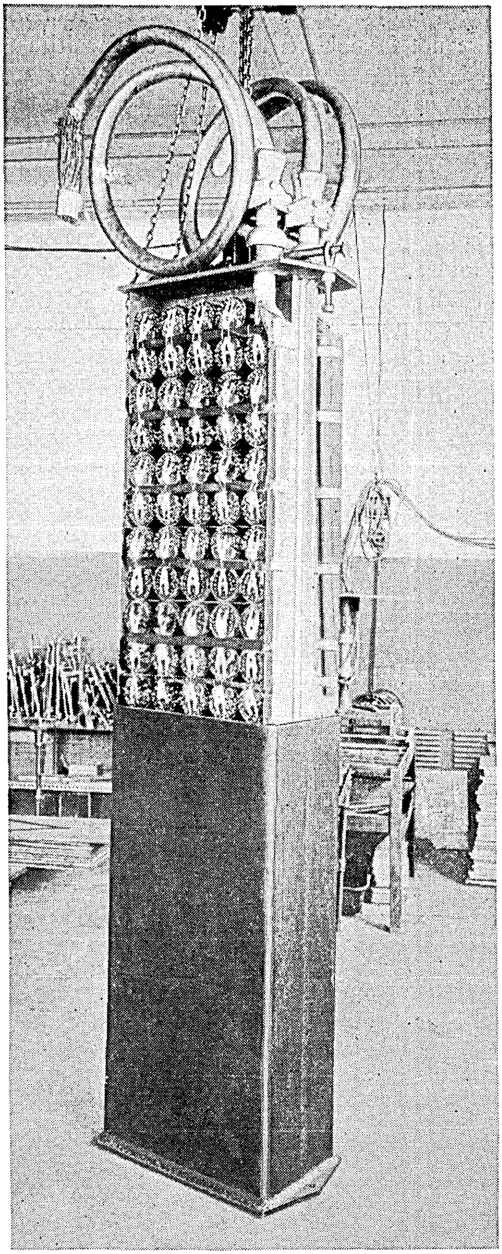


Fig. 11—Potting assembly—108 P-B type loading units in large welded steel case. Assembly ready for lowering into casing. Half of loading units are on far side of assembly frame, not visible in picture. The large number of loading coils makes necessary the use of 2 stub cables.

(11) Compressed Molybdenum-Permalloy Powder Core Loading Coils

11.1 *The Improved Core-Material*

The continuing search for still better core-materials culminated during the middle 1930's in the development of the 125-permeability compressed molybdenum-permalloy material, described in an A.I.E.E. paper[26] by V. E. Legg and F. J. Given, "Compressed Powdered Molybdenum-Permalloy for High-Quality Inductance Coils." In its intrinsic magnetic and electrical properties as used in voice-frequency loading coils, the improved permalloy-material is nearly as much superior to the old standard 75-permeability permalloy as this latter material was to the 35-permeability compressed powdered-iron which it superseded as standard during the late 1920's.

The improved permalloy owed its superior characteristics mainly to the inclusion of molybdenum in the alloy, the principal constituents of which are approximately 2% molybdenum, 81% nickel, and 17% iron. The molybdenum component substantially raises the permeability and the specific resistance, and materially reduces hysteresis. The eddy-current losses are much lower than in the 75-permeability permalloy material, because the effect of the higher permeability in increasing these losses is more than offset by the effect of the higher intrinsic resistance in reducing them.

In developing the material, it was considered to be very important to go as far as possible in improving the intrinsically favorable hysteresis properties. This was accomplished without adverse effects on permeability and eddy-current losses and other important properties, by annealing the pressed core rings in hydrogen at a much higher temperature than that previously used with material under treatment while exposed to the atmosphere. Also an improved type of particle insulation was developed. All in all, a large number of unusually difficult processing problems had to be solved in the research and development stages. Also, some of the processes required considerable additional attention during the manufacturing preparations for quantity production.

In the voice-frequency loading applications, the successful efforts to reduce hysteresis to a minimum fitted in with the preliminary economic design studies, which indicated it would be desirable to take advantage of the superior intrinsic properties of the new magnetic material by using it in smaller cores, so as to reduce loading-apparatus costs as much as possible without appreciably degrading transmission performance. In this connection, it is noteworthy that with any given magnetic-material the hysteresis losses become greater as the core size is reduced, in consequence of the greater intensity of magnetization.

The 125-permeability molybdenum-permalloy core material under discussion was developed primarily for voice-frequency uses. At much higher frequencies, lower permeabilities are necessary to prevent the core losses from becoming too high. Accordingly, other grades of compressed molybdenum-permalloy powder having lower permeability values are available. These are obtained by diluting the molybdenum-permalloy powder with inert material before pressing. Also, smaller-size particles are used. A new grade not described in the Legg-Given paper,[26] previously referred to, which has an effective permeability of 60, was used in small carrier loading coils for the Army spiral-four field cable during the war,[27] and is now being used in the cores of cable loading coils for 15-kc program transmission circuits which are described in Subdivision 13.3.

11.2 *M-type Molybdenum-Permalloy Core Loading Units*

The initial standard, molybdenum-permalloy core, phantom loading units which were coded in the M-series became available for commercial use during 1938. The individual coils were about 60% smaller than the standard, 75-permeability, permalloy-core coils which they superseded for use in new plant. In the headpiece, these size-relations are typified by coils D and C, respectively.

By design, the new coils had about the same d-c resistance and hysteresis loss as the superseded designs. Their eddy-current losses were considerably lower than those of the 75-permeability permalloy designs, and in consequence the total effective resistance was lower at the upper frequencies, thereby improving the steady-state frequency-distortion characteristics of the circuits in which they were used. In plant-design engineering, the new coils were accepted as being equivalent to the older coils. The residual inductance stability was a little better, and the telegraph-flutter distortion characteristics were considerably better. The susceptibility to superposed d-c magnetization, however, was worse. This minor impairment was the only adverse effect of the substantial increase in permeability, and the substantial reduction in coil size.

The development was timely in that the new coils were available for use in meeting the accelerating demand for toll cable loading that started in 1939 and continued for several years. In the five-year period 1938–1942, a total of about 800,000 side and phantom toll cable loading coils were manufactured for Bell System use notwithstanding the large installation of Type K carrier systems on non-loaded and unloaded cables that occurred in this period, thereby reducing the demand for additional, repeatered, loaded cable voice-frequency facilities.

The economic advantage of the broad-band carrier system is largely due to the fact that the cost of the conductors and the repeaters and of the distortion corrective-networks and regulating devices which are used to shape and control the transmission medium is shared by all of the transmission paths. Very valuable transmission advantages result from the relatively very high velocity of transmission over the non-loaded conductors.[o]

One of the effects of the new carrier systems' competition was to more than reverse the relative amounts of loading for new installations of 4-wire and 2-wire circuits—somewhat less than $\frac{1}{3}$ the total being provided for 4-wire circuits, during the period under consideration. The substantial cost-reductions, resulting from the introduction of the 75-permeability permalloy coils, of course materially limited the additional savings that could be realized by further size-reduction. Notwithstanding this, and taking into account also the declining demand for toll cable loading, the development of the M-type loading units turned out to be a very profitable operation, in terms of the reduced costs of new plant.

11.3 *SM and MF-type Molybdenum-Permalloy Core Loading Units*

These war-emergency designs owed their existence to the necessity for conserving strategic materials, especially nickel. They use half as much molybdenum-permalloy as the M-type loading units. The use of a new type of insulation, Formex enamel,[p] on the conductors in place of a combination of cotton with an older type of enamel, greatly improved the winding space-factor and minimized the increase in d-c resistance that necessarily followed from the 50% reduction in coil size. In the frontispiece, Coil E is one of these new coils, Coil D being the superseded coil.

To minimize delays in introducing the war-emergency designs into commercial use, which began during 1942, they were (initially) assembled in the unit-containers and potted in the loading coil cases that had been designed for the M-type loading units. They were initially coded as the SM-type loading units.

Subsequently, to conserve steel, entirely new unit-assembly arrangements were made in new smaller-size unit containers and new smaller-size loading coil cases were developed to take full advantage of the loading coil size-reduction. The coils themselves were not changed but new code-designations were assigned in the "MF" series, in conformity with the long established practice of coding phantom loading units in their unit containers.

The d-c resistances of the SM and MF loading units are about 25% higher than those of the corresponding M-type loading units, and the hysteresis effects are about 40% greater, under similar operating conditions. The other core-losses are unchanged in magnitude because they do not vary as

[o] Published information regarding the cable carrier systems is given in Bibliography References (28) and (29).

[p] This improved wire-insulation had already found valuable uses in new exchange area loading coils, as discussed in Section 22.

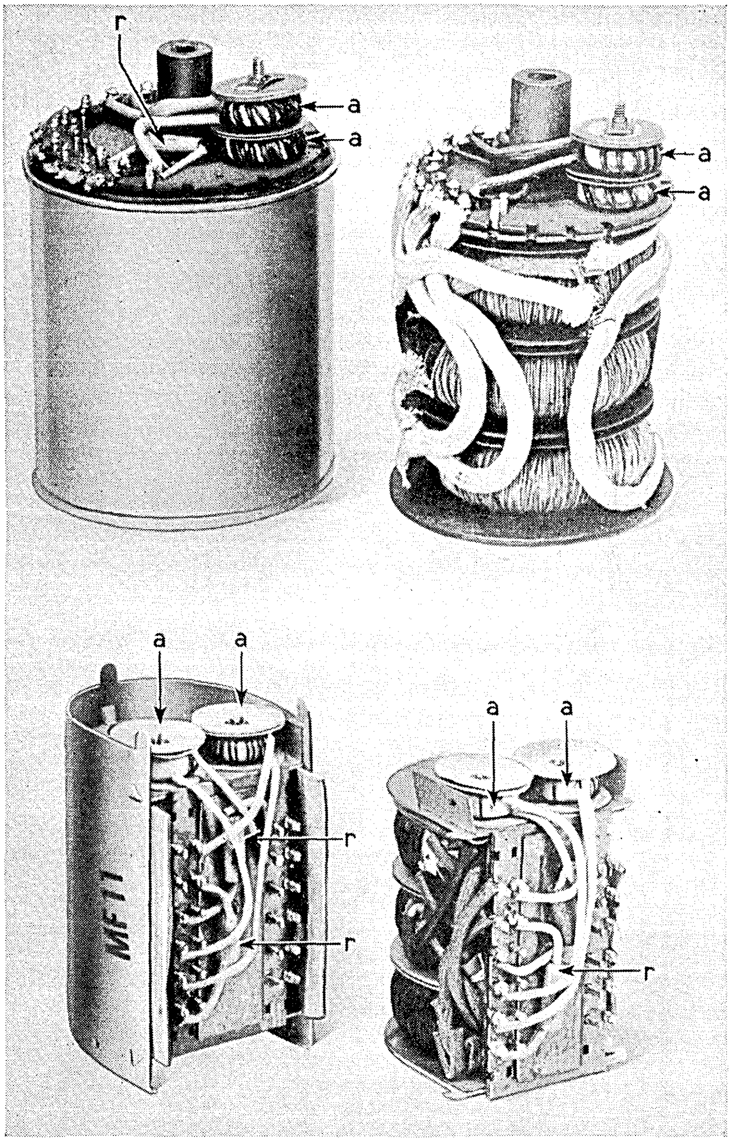


Fig. 12—M-type and MF type phantom loading units. Before and after placement in their unit shielding containers. M-type unit above the MF-type unit. "a" designates midget inductance coils (used in crosstalk adjustments); "r" designates midget resistors (used in crosstalk adjustments). The loading units shown outside the containers are not the same as those shown inside the containers.

a function of coil size. For the most important types of loading, the increases in attenuation that result from the higher coil resistances are in the range

1 to 3% at 1000 cycles on 19 ga. cables. At the top voice-frequencies, the attenuation increments are less than twice as large as the 1-kc increments. In 50-mile repeater sections, the resulting increases in the bare-line attenuation are of the order of 0.2 to 0.4 db at 1000 cycles depending on the type of loading and much of this increment-loss usually can be recovered by raising

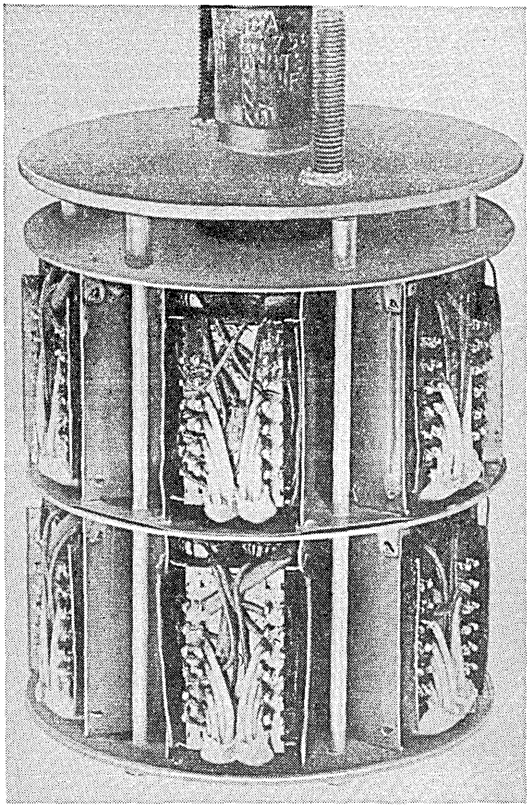


Fig. 13—Potting assembly methods—medium and large size complements of MF-type units. For placement in cylindrical, thin steel, casings.

the repeater gains. The stability of residual inductance is as good as that in the M-type loading units.

The impaired hysteresis effects, above mentioned, are accompanied by increased non-linear distortion, including telegraph-flutter effects. Also the smaller coils are somewhat more susceptible to changes in inductance and effective resistance when the circuits in which they are used have steady currents superposed upon them during talking intervals.

Additional information regarding the smallest loading units is included in an A.I.E.E. paper[30] by S. G. Hale, A. L. Quinlan, and J. E. Ranges, "Recent Improvements in Loading Apparatus For Telephone Cables."

The various relative transmission-impairments, above mentioned, were accepted in advance as being tolerable from the service standpoint under war conditions, considering the types of circuits required and their probable relatively short lengths. At that time it was thought possible that better loading units, not necessarily as good as the M-type units, might be warranted in the post-war period. Meanwhile, the service experience with the

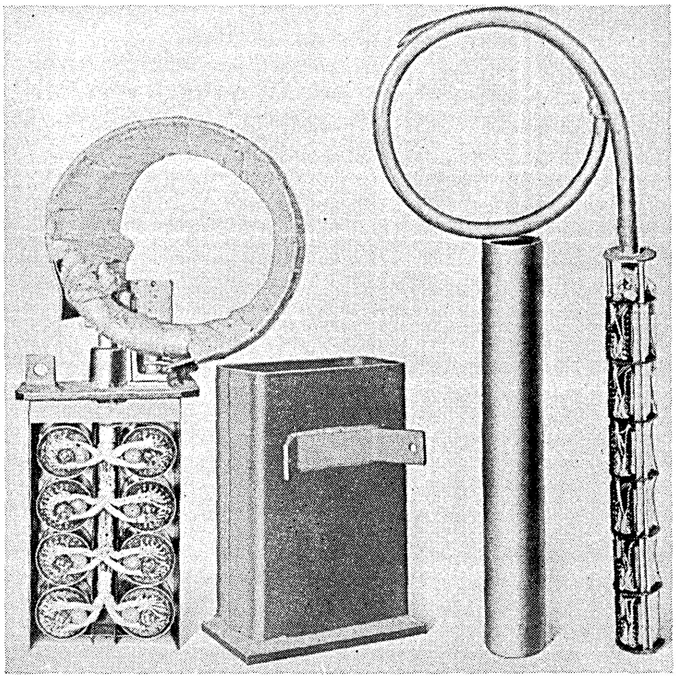


Fig. 14—Potting assembly methods—small complements of M-type and MF-type loading units. Note use of lead sleeve casing for MF type units, and relatively large welded-steel casing for M-type units.

small loading units has been reasonably satisfactory, largely because of the easing-up of performance requirements that is resulting from the restriction of toll cable loading to short-haul facilities. These seldom have more than two repeater sections, in consequence of the economic competition offered by present standard and proposed new, cheaper, carrier systems on non-loaded cable. Under these circumstances, the cost of developing better loading units which would probably increase the loading costs would be hard to prove in. Thus, the MF-type war-emergency loading units became post-war loading standards.

It should not be inferred from the foregoing that the MF (and SM) loading units have unimportant significance in the economy of Bell System toll cable plant-design. For example, during the period 1942–1949, a total of about two million side and phantom coils were manufactured for assembly in these smallest, standard, phantom loading units. A large majority was installed subsequent to VJ-day, as part of the Bell System program of plant expansion to meet the greatly increased demand for long-distance telephone service, and to restore the speed of service to the pre-war standards.

TABLE VI
ELECTRICAL DATA—MOST IMPORTANT VOICE-FREQUENCY PHANTOM LOADING UNITS

Loading Unit Code Nos.			Nominal Inductances (1)—mh		Approx. Avg. Resistances(2)—Ohms			
					Side Circuits		Phantoms	
			Side	Phantoms	d-c	1 kc	d-c	1 kc
P1, P1B			172.4	63.6	10.7	13.6	5.3	6.2
	M1		172.4	63.6	10.8	13.0	5.4	6.0
		MF1, SM1	172.4	63.6	13.8	16.4	6.9	7.6
P2, P2B			43.5	25.1	3.8	4.4	1.9	2.2
	M2		43.5	25.1	3.6	4.0	1.8	2.0
		MF2, SM2	43.5	25.1	4.5	5.0	2.2	2.5
P4, P4B			30.9	17.8	2.7	3.0	1.3	1.5
	M4		30.9	17.8	2.4	2.7	1.2	1.4
		MF4, SM4	30.9	17.8	3.1	3.4	1.5	1.7
P11B			88.7	50.2	6.1	7.0	3.05	3.5
	M11		88.7	50.2	6.3	7.2	3.1	3.6
		MF11, SM11	88.7	50.2	7.9	9.1	4.0	4.6

Notes: (1) The listed inductance values are the mean specification inductance values (at 1800 cycles).
(2) The coil resistance values allow for 19-gauge stub cables of 7.5-ft. external length.

(12) COMPARATIVE ELECTRICAL DATA, VOICE-FREQUENCY PHANTOM-GROUP LOADING UNITS

Comparative electrical data regarding the commercially most important, former standard and present standard, voice-frequency phantom-group loading units are given in Table VI, above. Those coded in the "P" and "PB" series used compressed permalloy-powder cores; the M, SM, and MF-type units used compressed, molybdenum-permalloy powder cores. Prior to the standardization of the P and PB-type units, the side circuit and phantom circuit coils had their own individual code numbers. They were interconnected in loading unit formation during the case assembly and cabling procedures.

The loading units having the digit 4 in their code designations are "full-weight" units for H31-18 voice-frequency entrance cable described in Part V of this review. The loading units having the digit 2 in their designations are used in H44-25 four-wire repeatered circuits. The other loading units are used in two-wire repeatered circuits, and also in circuits not long enough to require repeaters.

(13) Improved Loading for (Long-Distance) Cable Program Transmission Circuits

13.1 *General*

In the early days of radio chain-broadcasting (during the early 1920's), the links which connected the broadcasting stations with the studios where the programs originated usually were open-wire voice-frequency telephone circuits modified to meet the special requirements of this new type of service. In some instances, toll cable circuits were used for links not more than a few hundred miles long.

Where available, side circuits of H44-25 facilities, previously described, were preferred for the cable program-transmission service because of their high cut-off frequency. By using suitable equalizing and regulating networks in conjunction with modified one-way amplifiers, a fairly satisfactory transmission-medium could be obtained, providing a frequency-band ranging from about 100 cycles up to about 4000 cycles. While these circuits were adequate for speech broadcasting they were not equally satisfactory for classical music programs by symphony orchestras.

As early as 1924, some preliminary studies were started on entirely new types of loaded cable facilities primarily for use in transmitting the highest grade of radio-broadcast program material. These studies showed 16 ga. cable pairs to be desirable for long program circuits. At the time, however, there was considerable question as to what the ultimate performance-requirements should be, and some doubt as to whether the commercial demand would be sufficiently large to warrant the high cost of developing and providing a suitable new type of facility.

During the following years, there was a large increase in the number of broadcasting stations and in the need for inter-connecting links. Also the toll cable network expansion had commenced to accelerate rapidly. Beginning about 1926, the new toll cables that were installed usually included a small complement of 16-gauge non-quadded pairs (generally 6) in anticipation of their ultimate use for improved program facilities, if the then expected demand should eventually materialize. Studies of improved loading, and of the very important equalization and temperature regulation problems were renewed. This work progressed sufficiently so that early in 1927 a trial installation of H22 loading was planned for program facilities in a new cable between New York and Philadelphia. This loading had a 40% higher cut-off frequency than the H44 loading previously mentioned, and a 30% lower nominal impedance. It was expected that the program frequency-band could be stretched to a 6000-cycle top. It used a new loading coil described later on.

13.2 *B22 Program Facilities*

While the preparation for the H22 trial installations was still under way, further experimental and theoretical studies showed it would be desirable to provide an equalized transmission-band from about 50 cycles to about 8000 cycles, so as to obtain a margin for probable future improvements in broadcasting services. Accordingly, a decision was reached in September 1927 to develop a new type of cable program facility that would be satisfactory for lengths of 2000 miles or more. A new type of loading, designated B22 (22 mh loading coils at 3000-ft. spacing), was authorized and a trial installation on cable circuits looped back and forth between New York and Pittsburgh was planned.

The development of the new loading was only a small part of the total development effort. The equalizing and temperature-regulation problems were much more difficult than those previously encountered in the development of long-distance telephone message facilities, partly because of the much wider transmission-band and partly because of the more severe limits necessarily imposed. Improved repeaters were required and crosstalk and noise problems demanded very serious attention. A comprehensive account of the development work and a detailed description of the B22 cable program-facility is given in an A.I.E.E. paper[31] by Messrs. A. B. Clark and C. W. Green. Accordingly, the remaining discussion herein is limited to the loading.

The theoretical cut-off frequency of B22 loading is a little over 11,000 cycles; its nominal impedance is closely equal to that of H44 loading, previously mentioned. The theoretical nominal velocity of wave propagation is about 20,000 miles per second. The use of phase equalization networks of 8 kc band width, however, slows down the actual velocity to about 13,000 mi/sec. The attenuation on 16 ga. pairs at average temperature is about 0.24 db/mi at 1000 cycles, and ranges from about .14 db/mi at 35 cycles to about .38 db/mi at 8000 cycles.

The new 22 mh non-phantom type loading coils used compressed, permalloy-powder cores similar in size to those of the toll cable side circuit and phantom loading coils previously described in general terms. Non-linear distortion in the line was satisfactorily low in consequence of the favorable hysteresis characteristics of the permalloy-core coils.

The satisfactory completion of the system's trial-installation tests in 1929 resulted in a large amount of B22 loading being installed during 1930 and 1931 on 16-gauge pairs in cables that had been installed in advance of the facility development work, as previously mentioned, and in new cables.

During 1936, compressed, molybdenum-permalloy powder-core program loading coils became available for use in place of the permalloy-core coils above described. The new coils were smaller than their predecessors, and were substantially as good or better in the lower half of the working frequency band. At the high frequencies they were better than the permalloy-core coils, because of the lower eddy-current losses previously mentioned. An

Fig. 15—Molybdenum-permalloy core program circuit loading coil potted for installation within cable splice sleeves. The canvas bag shown in the picture is used to prevent the metal shielding-container of the coil from damaging the insulation of nearby conductors in the cable splice. The fabric strips are used to tie the coil case to the cable core at the splice.

important factor in the choice of coil size was to make the coil suitable for installation within the loading splice-sleeves so as to reduce potting and installation costs. This practice became quite general, especially at the "B" loading points intermediate between the "H" loading points where phantom loading units were installed on telephone message circuits in the same cable. In these "splice installations," the per-coil savings in potting and installation cost were large relative to the direct savings in loading coil costs resulting from size reduction. In order to realize these substantial combined savings as quickly as possible, the new program coils were given priority in the commercial exploitation of the molybdenum-permalloy core-material.

As an indication of the commercial importance of the program circuit loading above described, it is of interest to note that enough program coils were manufactured during the period 1928–1948 to provide loading for more than 100,000 pair-miles of B22 facilities. The demand for the new B22 loading is rapidly declining however, in consequence of the post-war development of program transmission facilities using combinations of channels in non-loaded cable carrier systems. From now on, it seems probable that new B22 loading will largely be limited to short extensions of existing loaded facilities, and to short links between cable carrier-system terminals and the points where broadcasting programs are picked up or delivered.

13.3 *15-kc Program Transmission Circuits*

The development of new loading for use on cable facilities transmitting FM broadcasting material was completed early in 1948. It provides a transmission band extending up to 15 kc and is suitable for use on circuits that connect FM broadcasting transmitters with their studios, in situations where its use will reduce costs as compared with the use of intermediate amplifiers.

Two weights of 15-kc loading are available, one having a nominal impedance equal to that of B22 loading (800 ohms), and a lighter-weight loading having a nominal impedance of about 480 ohms.

The 800-ohm loading generally uses 11 mh coils at spacings which give a theoretical cut-off frequency of about 23 kc, and a nominal velocity of about 20,000 miles per second. In 0.062 mf/mi cables the coil spacing is 1500 ft. and in "high-capacitance" exchange cables, it is 1100 ft., in order to obtain closely the same total loading section capacitance and similar values of nominal impedance and cut-off frequency. Since studio-transmitter circuits usually include tandem combinations of component cables which have appreciably different unit-length mutual capacitances, this loading plan minimizes reflection effects which would otherwise result from impedance differences at the junctions of the component cables.

800-ohm loading is sometimes provided on program pairs in coaxial cables by using 7.5 mh coils at 1000 ft. spacing. When the coaxial cables are installed in 1000 ft. lengths, this loading arrangement avoids the need for making the (expensive) extra cable splices that would be required if 1500 ft. spacing should be used. Although 50% more program loading coils are required, the small additional cost of the loading is negligible in relation to the savings in cable splicing costs. This 7.5 mh, 800-ohm loading has a nominal velocity of about 20,000 mi/sec, and a theoretical cut-off frequency of about 34 kc.

The 480-ohm loading uses 7.5 mh coils at spacings twice as long as those for the 800-ohm loading using 11 mh coils. The theoretical cut-off is a little over 19 kc and the nominal transmission velocity is about 36,000 miles per second.

The 800-ohm loading is superior in all transmission features to the 480-

ohm loading, especially at low frequencies. It is an acceptable substitute for B22 loading for transmission of AM broadcasting material, being appreciably better in the frequency range 4 to 8 kc, because of its much higher cut-off frequency, and it is substantially as good at low frequencies. Accordingly, to provide plant flexibility for possible future broadcasting service requirements, this 800-ohm 15-kc loading is being installed on short-haul program pairs in new toll cables in situations where the anticipated initial needs are for AM broadcast programs. When the loading is installed in the course of the cable installation, the slightly higher cost of the broader-band loading is negligible in comparison with the plant flexibility-advantage above cited.

The field of use for the 480-ohm 15-kc program loading is in short repeater sections where the transmission requirements permit its use, and where it is desirable to take advantage of its lower cost, relative to the 800-ohm loading. This cost advantage is maximum when the loading has to be applied to cables previously installed.

The loading coils for the new 15-kc program loading are of the same size as the molybdenum-permalloy core coils developed for B22 loading. Accordingly, when installation conditions are favorable, economies can be achieved by placing the coils within the loading splices. The new program loading coils use 60-permeability, compressed, molybdenum-permalloy powder cores. This relatively new grade of molybdenum-permalloy has much lower eddy-current losses than the 125-permeability grade used in voice-frequency loading coils (and in B22 program loading), and is very advantageous in the reduction of attenuation over the upper two-thirds of the 15-kc transmission-band. The 60-permeability cores also have superior non-linear distortion-characteristics.

(14) Control of Crosstalk

The control of crosstalk between adjacent telephone circuits has been a very important problem from the beginning of the use of loading, especially in loaded cables, and has been most difficult in long loaded, repeatered, quadded toll cables.

14.1 *Cable Unbalances*

The effect of loading in increasing the circuit impedance raises the voltages which act upon the unavoidable cable capacitance-unbalances, and reduces the magnitudes of the line currents which act upon the series resistance (and inductance) unbalances. Consequently, with the introduction of loading, it became necessary to improve the design of the cables so as to materially reduce the capacitance unbalances. This was accomplished by using different lengths of twist in adjacent pairs and in adjacent layers, and by using different lengths of lay in adjacent layers.

The development of quadded cable and of phantom-group loading during the period 1908–1910 introduced especially difficult requirements in the control of crosstalk between the phantom circuits and their side circuits, and to a lesser degree between the associated side circuits. Notwithstanding further improvements in design and great care in manufacture, it became necessary during the installation of the quadded cables to resort to the use of special splicing procedures which reduced the total capacitance-unbalance per loading section to satisfactory values.

For nearly a decade it was the practice in the control of phantom-to-side and side-to-side crosstalk to make the capacitance-unbalance test-splices at seven approximately even-spaced intermediate splicing-points in each loading section. By about 1919, further improvements in design and in manufacturing processes made it feasible to reduce the number of test splices per loading section to 3. This practice is still general for quadded cables used principally for two-wire repeatered facilities. In the period of very rapid extension of the installation of long repeatered, loaded four-wire circuits (1929–1931), it was found feasible to limit the number of test splices per loading section to one, by using supplementary balancing-condensers to reduce the high residual capacitance-unbalances on two-wire circuits, and by using additional balancing-condensers at one end of each repeater section on four-wire circuits, when required to obtain satisfactory low over-all crosstalk. The one-way transmission in the individual, oppositely-bound, paths of the four-wire circuits was a basic factor in making feasible these field-adjustment simplifications, thereby reducing installation costs.

14.2 *Loading Coil Crosstalk*

Although the side circuit and phantom loading coils never had objectionable design unbalances, it has always been difficult in manufacture to realize the inherent symmetry of their design. The series inductance unbalances have been the most troublesome accidental unbalances. In the early days, reasonably satisfactory control of crosstalk between circuits in the same phantom-group was obtained by care in manufacture, and by adjusting the inductance unbalances to the nearest turn.

A general, substantial, tightening of the crosstalk limits became necessary when repeaters came into general use, because of the effect of the repeaters in amplifying the unwanted crosstalk along with the wanted conversation, and because of the great increases in circuit length that the repeaters made feasible. During a period in the early 1920's, when intensive development was under way to reduce the loading apparatus crosstalk, the over-all repeater-section crosstalk was controlled by poling the unbalances of the coils against the cable unbalances in the loading sections near the repeaters.

The new development work, above referred to, included greater refinements in the manufacturing processes, including greater accuracy in the subdivision of the core winding-space, more uniform winding-arrangements in layer formation by automatic winding-machines, and the use of external series-inductance adjustments by means of preselected, miniature, inductance coils, which resulted in much smaller residual unbalances than could be obtained by the nearest full-turn adjustments. More precise crosstalk measurement-circuits were developed, and the practice started of making the crosstalk adjustments and measurements in test circuits having, at the test frequency, impedance characteristics approximating those of the lines in which the coils under test are used. (Previously a relatively insensitive test-circuit having "compromise" impedance characteristics had been used for all types of side and phantom coils.) Also, the factory crosstalk adjustments were concentrated on minimizing near-end crosstalk in loading units intended for two-wire circuits, since in such circuits the far-end crosstalk is relatively unimportant. In loading apparatus intended for use on four-wire circuits, the crosstalk adjustments were made to minimize far-end crosstalk because the associated phantom-group circuits are always used in the same direction of transmission and the one-way repeaters associated with them block the propagation of near-end crosstalk. The external inductance adjustments, above referred to, were especially designed for the reduction of phantom-to-side crosstalk. Other adjustments and assembly processes controlled crosstalk between associated side circuits.

As a result of the various improvements above referred to, the average phantom-to-side crosstalk in the loading apparatus was reduced to about one-fourth of the earlier values. These reduced values were of the same order as the cable crosstalk in the individual loading sections after the completion of the capacitance-unbalance test-splicing.

In the late 1920's and early 1930's, when the need for improved transmission systems on two-wire repeatered circuits resulted in the development of the H88-50 and B88-50 facilities previously described, another development campaign was started to improve loading apparatus crosstalk-performance. This resulted in the reduction of the average loading coil crosstalk to values much smaller than those caused by the cable residual capacitance-unbalances after test splicing, so that the effective repeater-section crosstalk is not significantly larger than that which would result if it were possible to manufacture perfectly balanced loading coils. Important factors in this improved performance were the use of series, external, resistance adjustments in particular loading units where worth-while crosstalk reduction could be thus obtained, and the use of a new series of inductance elements having closer gradations in their inductance values to more closely approach the theoretical optimum adjustments. More compact assembly-arrangements of the phantom loading units in individual shielding containers also contributed to the more favorable crosstalk-performance, along with other process improvements. (Refer to Fig. 10, page 186.)

The improved crosstalk-performance above discussed, which was achieved in the development of the PB-type permalloy-core phantom loading units, previously described, was maintained in the commercial production of the smaller-size M-type molybdenum-permalloy core loading units which superseded the PB-type loading units, and in the present standard MF-type loading units. (Refer to Fig. 12, page 192.)

It seems improbable that additional improvements in loading unit crosstalk-performance will be needed in the future, in view of the trend during recent years of restricting voice-frequency phantom-group loading to short-haul facilities.

A more comprehensive account of the successful struggle to control cable crosstalk in loaded toll cables is given in a 1935 Bell Telephone Quarterly article[32] by Mr. M. A. Weaver. An A.I.E.E. paper,[8] previously referred to, includes considerable additional information on the crosstalk-reduction work on the loading coils, up to 1926.

Bibliography

1. George A. Campbell, "Collected Papers," A.T.&T.Co., New York, 1937. Introduction by E H. Colpitts; see also "Three Early Memoranda on Loading," page 10.
2. M. I. Pupin, "Wave Transmission over Non-Uniform Cables and Long Distance Air Lines," *Trans. A.I.E.E.*, Vol. XVII, 1900, p. 445. Also, Pupin *U. S. Patent Nos.* 652,230 and 652,231, June 19, 1900.
3. George A. Campbell, "On Loaded Lines in Telephonic Transmission"—*Philosophical Magazine*, March 1903 (reprinted in Reference No. 1, based entirely on 1899 theoretical studies and experiments).
4. Vaschy, La Lumiere Electrique, Jan. 12, 1889.
5. O. Heaviside, "Electromagnetic Theory," Vol. 1, 1893, p. 441.
6. B. Gherardi, "Commercial Loading of Telephone Circuits in the Bell System," *Trans. A.I.E.E.*, Vol. XXX, 1911, p. 1743.
7. Roger B. Hill and Thomas Shaw, "Hammond V. Hayes: 1860–1947," *Bell Telephone Magazine*, Vol. XXVI, Autumn 1947, pp. 150–173.
8. Thomas Shaw and William Fondiller, "Development and Application of Loading for Telephone Circuits," *Trans. A.I.E.E.*, Vol. XLV, Published in *Bell System Technical Journal*, Vol. V, April 1926, pp. 221–281.
9. Thomas Shaw, "The Conquest of Distance by Wire Telephony," *Bell System Technical Journal*, Vol. XXIII, October 1944.
10. W. Fondiller and W. H. Martin, "Hysteresis Effects with Varying Superposed Magnetizing Forces," *Trans. A.I.E.E.*, Vol. XL, 1921.
11. B. Gherardi and F. B. Jewett, "Telephone Repeaters," *Trans. A.I.E.E.*, Vol. XXXVIII, Nov. 1919.
12. R. S. Hoyt, "Impedance of Loaded Lines and Design of Simulating and Compensating Networks," *B.S.T.J.* April 1923.
13. Buckner Speed and G. W. Elmen, "Magnetic Properties of Compressed Powdered Iron," *Trans. A.I.E.E.*, Vol. XL, 1921, p. 596.
14. A. B. Clark, "Telephone Transmission Over Long Cable Circuits," *Trans. A.I.E.E.*, Jan. 1923; *B.S.T.J.*, Jan. 1923; *Electrical Communication*, Feb. 1923.
15. H. S. Osborne, "Telephone Transmission Over Long Distances," *Trans. A.I.E.E.*, Oct. 1923; *Electrical Communication*, Oct. 1923.
16. A. B. Clark and R. C. Mathes, "Echo Suppressors for Long Telephone Circuits," *Trans. A.I.E.E.*, June 1925, p. 618.
17. G. A. Campbell, "Physical Theory of the Electric Wave Filter," *B.S.T.J.* Vol. 1, Nov. 1922.

18. J. R. Carson, "Theory of the Transient Oscillations of Electrical Networks and Transmission Systems," *Trans. A.I.E.E.*, Vol. XXXVIII, 1919, p. 345.
19. J. H. Bell, R. B. Shanck and D. E. Branson, "Metallic Polar-Duplex Telegraph System for Cables," *Trans. A.I.E.E.*, Vol. XLIV, 1925.
20. B. P. Hamilton, H. Nyquist, M. B. Long and W. A. Phelps, "Voice-Frequency Carrier Telegraph Systems for Cables", *Trans. A.I.E.E.*, Vol. XLIV, 1925.
21. A. L. Matte, "Advances in Carrier Telegraph Transmission," *B.S.T.J.*, April, 1940.
22. R. B. Shanck, "Neutralization of Telegraph Cross Fire," *B.S.T.J.*, July, 1926.
23. H. D. Arnold and G. W. Elmen, "Permalloy, An Alloy of Remarkable Magnetic Properties," *Journal of the Franklin Institute*, Vol. 195, 1923.
24. W. J. Shackelton and I. G. Barber, "Compressed Powdered Permalloy, Manufacture and Magnetic Properties," *Trans. A.I.E.E.*, Vol. 47, 1928, p. 429.
25. W. H. Martin, "Transmitted Frequency Range for Telephone Message Circuits," *B.S.T.J.*, Vol. IX, July 1930.
26. V. E. Legg and F. J. Given, "Compressed Powdered Molybdenum-Permalloy for High-Quality Inductance Coils," *Bell System Technical Journal*, Vol. XIX, 1940, p. 385.
27. J. E. Ranges, "Loading The Spiral-4 for War," *Bell Labs. Record*, Oct. 1946.
28. A. B. Clark and B. W. Kendall, "Carrier in Cable," *B.S.T.J.*, July 1933.
29. C. W. Green and E. I. Green, "A Carrier Telephone System for Toll Cables," *B.S.T.J.*, Jan. 1938.
30. S. G. Hale, A. L. Quinlan and J. E. Ranges, "Recent Improvements in Loading Apparatus for Telephone Cables," *Trans. A.I.E.E.*, Vol. 67, 1948.
31. A. B. Clark and C. W. Green, "Long Distance Cable Circuits for Program Transmission," *Trans. A.I.E.E.*, Vol. 49, 1930, p. 1514.
32. M. A. Weaver, "The Long Struggle Against Cable Crosstalk," Bell Telephone Quarterly, Jan. 1935.

Technical Publications by Bell System Authors Other Than in The Bell System Technical Journal

*Generalizations of the Weiss Molecular Field Theory of Antiferromagnetism.** P. W. ANDERSON.[1] *Phys. Rev.*, v. 79, pp. 705–710, Aug. 15, 1950.

ABSTRACT—A Weiss field calculation has been carried out for antiferromagnetism in more complicated structures than the usual calculation allows and has been shown to give results more detailed and more consistent with experimental evidence on the magnetic properties of such structures than does the simpler theory.

Atomic Moments of Ferromagnetic Alloys. R. M. BOZORTH.[1] Letter to the editor. References. *Phys. Rev.*, v. 79, p. 887, Sept. 1, 1950.

Domain Structure of a Cobalt-Nickel Crystal. R. M. BOZORTH[1] and J. G. WALKER.[1] Letter to the editor. *Phys. Rev.*, v. 79, p. 888, Sept. 1, 1950.

*Recording Fluxmeter of High Accuracy and Sensitivity.** P. P. CIOFFI.[1] *Rev. Sci. Instruments*, v. 21, pp. 624–628, July 1950.

ABSTRACT—A recording fluxmeter has been developed which employs one or two integrators and a double element L and N Speedomax recorder for tracing magnetization curves directly on standard coordinate paper. The response of the recorder pen drive mechanism is proportional to the flux density, B, and is controlled by the B integrator. The response of the paper drive mechanism is proportional to the magnetizing force, H, and is controlled either by the magnetizing current, when the specimen is in the form of a ring, or by the H integrator when the specimen is in the form of a bar. Ayrton shunt networks provide flexible B and H scale adjustments. High accuracy and sensitivity are obtained by minimizing the causes of drift. At maximum sensitivity, four interlinkages give a deflection of one mm.

*Young's Modulus and Its Temperature Dependence in 36 to 52 pct Nickel-Iron Alloys.** M. E. FINE[1] and W. C. ELLIS.[1] *Jl. Metals*, v. 188, pp. 1120–1125, Sept. 1950.

ABSTRACT—Young's modulus and its temperature coefficient in 36 to 52 pct Ni-Fe alloys depend upon composition and also the straining-annealing history. Alloys near 42.5 pct Ni, when worked cold and then annealed at 400° or 600°C, have nearly zero mean thermoelastic coefficients between −50° and 100°C. A discussion of the theory is given.

* A reprint of this article may be obtained on request to the editor of the B.S.T.J.
[1] B.T.L.

Ultra-High Vacuum Ionization Manometer. J. J. LANDER.[1] *Rev. Sci. Instruments*, v. 21, pp. 672–673, July 1950.

ABSTRACT—This note describes an ionization menometer which indicates pressures more than two decades below the lower limit usually encountered at about 1×10^{-8} mm of Hg with other manometers. This limit depends on the design of the gauge; however, values reported for various gauges do not differ much from that given above. Commonly the limit is observed as a lowest reading obtained despite recourse to more or less drastic methods of producing lower pressures, or a variety of changes in gauge design. The flash filament method of pressure measurement has been used to measure lower pressures and at the same time to indicate the lower limit of an ionization guage.

*Metallized Paper for Capacitors.** D. A. MCLEAN.[1] *I.R.E., Proc.*, v. 38, pp. 1010–1014, Sept., 1950.

ABSTRACT—Metallized capacitor paper is attracting widespread interest as a way of reducing capacitor size. In metallized paper capacitors, the usual metal foil is replaced by a thin layer of metal evaporated onto the surface of the paper. Lacquering the paper prior to metallizing increases the dielectric strength and insulation resistance, reduces atmospheric corrosion of the metal, and diminishes the rate of loss of electrode metal by electrolysis. Owing to the extreme thinness of the metal layer, metallized paper capacitors are subject to a type of failure not ordinarily found in conventional capacitors. This type of failure consists of the loss of electrode by electrolysis and occurs under d-c. potential when the ionic conductivity is high, as results, for example, from the presence of moisture. For this reason, it is recommended that special precautions be taken to keep the ionic conductivity low, in particular with respect to thorough and effective drying and sealing of the capacitor units.

Thermoelastic Stress Around a Cylindrical Inclusion of Elliptic Cross Section. R. D. MINDLIN[1] and H. L. COOPER.[1] *Jl. Applied Mech.*, v. 17, pp. 265–268, Sept. 1950.

ABSTRACT—The two-dimensional equations of thermoelasticity are solved for the case of a uniform temperature change of an infinite medium containing a cylindrical inclusion of elliptic cross section. The problem is treated as one of plane strain in elliptic co-ordinates, and the solution is given in closed form. Formulas and curves are given for the maximum values of various components of stress at the interface between the inclusion and the surrounding medium.

*Magnetostriction of Permanent Magnet Alloys.** E. A. NESBITT.[1] Bibliography. *Jl. Applied Phys.*, v. 21, pp. 879–889, Sept. 1950.

* A reprint of this article may be obtained on request to the editor of the B.S.T.J.
[1] B.T.L.

ABSTRACT—In order to obtain a better understanding of the mechanism of coercive force in modern permanent magnets, magnetostriction measurements have been made on various alloys having coercive forces from 50 to 600 oersteds. The results can be summarized by discussing two types of alloys. First are the older carbon-hardening permanent magnets, and for these alloys high coercive force and high magnetostriction occur together. Second are the new carbon-free permanent magnets and for these alloys high coercive force does not occur with high magnetostriction. In fact for the Mishima alloys having compositions near 29 per cent nickel, 12.5 per cent aluminum, and 58.5 per cent iron, cooled at the rate of 3°C per second (coercive force 400 oersteds), the magnetostriction actually passes through zero. This is contrary to the classical strain theory of coercive force which states that the latter is proportional to the product of the magnetostriction and internal stress. To explain the mechanism of coercive force for these alloys it is necessary to resort to more recent theories.

*Stress Analysis for Compressible Viscoelastic Materials.** W. T. READ, JR.[1] *Jl. Applied Phys.*, v. 21, pp. 671–674, July 1950.

ABSTRACT—Mathematical methods of stress analysis are presented for linear, compressible, viscoelastic, or anelastic, materials such as metals at high temperatures or high polymers with small strains. For such materials stress, strain and their time derivatives of all orders are related by linear equations with coefficients which are material constants. Fourier integral methods are used to show that static elasticity solutions can be used to determine the time dependent stresses in viscoelastic bodies with any form of boundary conditions.

If stress and double refraction and their time derivatives are also linearly related, the standard photoelastic techniques can be used to determine the directions and difference in magnitude of the time dependent principal stresses, even though the principal stress axes do not coincide with the polarizing axes and both vary with time. When viscoelastic models are used in photoelastic studies, the time variation of the stress distribution in the model represents a first approximation to the dependence of the stress in the elastic prototype on Poisson's ratio.

*Metallurgy Behind the Decimal Point.** E. E. SCHUMACHER.[1] *Jl. Metals*, v. 188, pp. 1097–1110, Sept. 1950.

ABSTRACTS—No one property has a monopoly as to being disproportionately affected by minor elements. Nearly all properties are affected, but there is time here to include only a selected few. I have chosen, therefore, three of general interest: strength, magnetic, and electrical properties. I shall inquire into both the mechanisms and consequences of these dispro-

* A reprint of this article may be obtained on request to the editor of the B.S.T.J.
[1] B.T.L.

portionate effects and try to share with you the fascination and challenge of this field.

Chain Feed Carries Springs Through Forming Die. J. D. THOMPSON[2] and G. W. RADA.[2] *Am. Mach.*, v. 94, pp. 86–87, Oct. 2, 1950.

Punch-Press Tooling Works from Cams, Makes Cams. J. H. TOMLIN.[2] *Am. Mach.*, v. 94, pp. 106–108, Sept. 18, 1950.

ABSTRACT—Literally millions of cam-shape combinations are possible in the telephone-exchange contact cams made for Bell Telephone; yet only two punches and two dies do all the work . . . Two cams are cut at a time from same-size master cams in a punch-press setup that allows a switch to a new cam shape in a matter of seconds.

*Construction of Cold-Cathode Counting or Stepping Tubes.** M. A. TOWNSEND.[1] *Elec. Engg.*, v. 69, pp. 810–813, Sept. 1950.

ABSTRACT—Electronic digital counters are capable of performing at high speeds many of the functions which are performed at low speeds by chains of relays and mechanical stepping switches. Here is described a new principle of tube construction by means of which the position of a glow discharge can be made to step along a row of cold cathodes under the control of input pulses.

*Ferromagnetic Domains.** H. J. WILLIAMS.[1] References. *Elec. Engg.*, v. 69, pp. 817–822, Sept., 1950.

ABSTRACT—Ferromagnetism is based on the property of domains. These are tiny regions within a magnetic substance. They have this special characteristic: most of the elementary atomic magnets contained in a particular domain have their spins oriented in the same direction. Domain sizes and shapes are the result of an attempt of the ferromagnetic system to reach a state that minimizes the magnetostatic, magnetostrictive, domain-boundary, and other energies.

416A-Tube for Microwave Relays. K. P. DOWELL.[2] *FM*, v. 10, pp. 20–22, Aug., 1950.

ABSTRACT—For construction of its cross-country microwave relay chain, the Bell System required a 4,000-mc. amplifier tube having a greater gain-bandwidth product than any available at the time. The 416A planar triode, shown in Fig. 1, was developed especially for this application. Because of its exceptionally close interelectrode spacing, new techniques were necessary for factory production of the tube. It is the purpose of this paper to show some of the unique operations employed in its manufacture and assembly.

*On the Acoustics of Coupled Rooms.** C. M. HARRIS[1] and H. FESHBACH. *Acoustical Soc. Am., Jl.*, v. 22, pp. 572–578, Sept. 1950.

* A reprint of this article may be obtained on request to the editor of the B.S.T.J.
[1] B.T.L.
[2] W.E.Co.

Abstract—In many acoustically coupled systems the methods of geometrical acoustics do not apply. Reverberation formulas as ordinarily used would lead to incorrect results. This paper approaches the problem of coupled rooms from the "wave" point of view, treating the coupled rooms as a boundary value problem in obtaining an approximate solution. The results explain some discrepancies noted by earlier researchers between experiment and predictions from geometrical acoustics; for example, the dependence of absorption in a room on the position of the open area which couples the room to an adjacent one. For the case where the window area which couples one room to another is comparable in size with the partition which separates the rooms, the effect of the partition will be least when it is at a particle-velocity node. For the case where the window area is small compared with the partition which separates the two rooms, the effect of the coupling window depends on the square of the unperturbed pressure at the window. Thus the effect of the window varies with position and is least at a pressure node. Experimental data on isolated modes of vibration of a coupled system are given which check the results predicted by this application of the wave theory.

*Quantitative Spectrochemical Analysis of Ashes, Deposits, Liquids, and Miscellaneous Samples.** E. K. Jaycox.[1] References. *Anal. Chem.*, v. 22, pp. 1115–1118, Sept. 1950.

Abstract—A general technique is described which is applicable to the quantitative spectrochemical analysis of a wide variety of materials. Sample preparation, the incorporation of spectrochemical buffers, and excitation procedures are discussed for typical cases which illustrate the scope and possibilities of the method. Examples include the analysis of the ashes of rubber, plastics, paper, and cloth; deposits on walls of vacuum tubes and other surfaces; water, oils, and other liquids; and miscellaneous solid materials.

*Response Peaks in Finite Horns.** C. T. Molloy.[1] *Acoustical Soc. Am., Jl.*, v. 22, pp. 551–557, Sept. 1950.

Abstract—It is the purpose of this paper to discuss the theory of the axial response curves of the class of acoustical systems comprising a driving unit and a finite "Hyperbolic Horn". The term "Hyperbolic Horn" as used in this paper denotes horns of the type first discussed by Salmon. It is proposed to derive an expression from which the response curve may be calculated. A method will also be given by means of which the peaks in this response curve may be located without the necessity of computing the whole curve. Finally the converse problem will be discussed, namely, how to

* A reprint of this article may be obtained on request to the editor of the B.S.T.J.
[1] B.T.L.

choose the driving unit and horn parameters so that one of the response peaks will occur at a predetermined frequency.

Hard Rubber. H. PETERS.[1] Bibliography. *Ind. & Engg. Chem.*, v. 42, pp. 2007–2019, Oct. 1950.

ABSTRACT—The chemical engineer who picks hard rubber for various special uses usually does so on the basis of its availability, its relatively reasonable cost, its good physical and chemical properties, its ease of machining and fabrication, and its excellent resistance to a great variety of chemicals. This annual review of 1949, like the previous one in 1948 (32), deals with the above in addition to some fundamental studies on natural and synthetic hard rubbers. Heretofore, these basic investigations were usually centered around hard rubber prepared from natural rubber; now there is a decided interest in hard rubber prepared from the synthetics. While admittedly this interest is at a very low ebb, the trend is definitely upward and probably will continue that way for years to come.

*Holes and Electrons.** W. SHOCKLEY.[1] *Physics Today*, v. 3, pp. 16–24, Oct. 1950.

ABSTRACT—Some new experiments in transistor electronics are described here in which concepts suggested by theory have been verified directly by experiment.

*Metallized Paper Capacitors.** J. R. WEEKS.[1] *I.R.E., Proc.*, v. 38, pp. 1015–1018, Sept., 1950.

ABSTRACT—Metallized paper capacitors are being introduced into telephone apparatus wherever size is of prime importance. It is shown that low-voltage metallized paper capacitors with about half the volume of a foil-paper capacitor of conventional design have about the same characteristics as the latter. Performance data are discussed which indicate that such capacitors will give long service when used within their voltage rating and when well-protected against moisture. It is pointed out that this type of capacitor should be used within its voltage rating if sparking with its attendant circuit noise is to be avoided. When sparking does occur due to abnormal voltage conditions no permanent damage results.

* A reprint of this article may be obtained on request to the editor of the B.S.T.J.
[1] B.T.L.

Contributors to this Issue

R. S. CARUTHERS, B.S., University of Maryland, 1926; E.E., 1930; M.S., Massachusetts Institute of Technology, 1928. General Electric Company, 1926–28; U. S. Bureau of Standards, 1928–29. Bell Telephone Laboratories, 1929–. Prior to the war Mr. Caruthers was engaged chiefly in the development of the C, J and K Carrier Telephone Systems. During the war he was engaged in the design of radar equipment. Since the war his principal activities have been as project engineer for the N1 and O Carrier Telephone development.

J. J. GILBERT, A.B., University of Pennsylvania, 1909; Harvard, 1910–11; University of Chicago, 1911–12; E.E., Armour Institute, 1915; Instructor of Electrical Engineering, Armour, 1912–17. Captain, Signal Corps, 1917–19. Bell Telephone Laboratories, 1919–. Mr. Gilbert's work with the Laboratories has been concerned with submarine cable problems.

WILLIAM M. GOODALL, B.S., California Institute of Technology, 1928. Bell Telephone Laboratories, 1928–. Mr. Goodall has worked on research problems in connection with the ionosphere, radio transmission and early radio relay studies, radar modulators, microwave radio relay systems, and Pulse Code Modulation.

J. L. MERRILL, JR., Pennsylvania State College, B.S. in Electrochemistry, 1928; Elliot Research Fellow, 1928–30; M.S., 1930. American Telephone and Telegraph Company, Department of Development and Research, 1930–34; Bell Telephone Laboratories, 1934–. With the American Telephone and Telegraph Company Mr. Merrill worked on transmission problems related to the exchange area. After coming to the Laboratories he continued to work on exchange area transmission projects such as the transmission features of the Time and Weather Announcement Systems, Service Observing, Operator Training and like systems. During the war he was engaged with a group planning system operation of air raid warning and of tactical wire and radio networks. Since the war his efforts have been directed toward the design and application of repeaters for the exchange area plant.

RUSSELL C. MINER, B.S. in Electrical Engineering, University of California, 1929. Bell Telephone Laboratories, 1929–. Mr. Miner has been engaged in work chiefly concerned with the development of acoustical instruments.

EDWARD E. MOTT, B.S. and M.S. in Electrical Engineering, Massachusetts Institute of Technology, 1928. General Electric Company, 1926–28. Bell Telephone Laboratories, 1928–. Mr. Mott has been engaged in telephone instruments research and development, particularly in connection with various types of telephone receivers and related devices. During the war he was engaged in underwater sound studies.

R. L. PEEK, JR., Columbia College, A.B. 1921; School of Mines, Columbia University, Met. E., 1923. Bell Telephone Laboratories, 1924–. From 1924 to 1936 Mr. Peek was engaged in studies of materials and materials testing. Since 1936 he has been engaged in the development of electromagnetic switching apparatus and (in 1941 to 1945) other applications of magnetics and magnetostriction.

CLAUDE E. SHANNON, B.S. in Electrical Engineering, University of Michigan, 1936; S.M. in Electrical Engineering and Ph.D. in Mathematics, M.I.T., 1940. National Research Fellow, 1940. Bell Telephone Laboratories, 1941–. Dr. Shannon has been engaged in mathematical research principally in the use of Boolean Algebra in switching, the theory of communication, and cryptography.

THOMAS SHAW, S.B., Massachusetts Institute of Technology, 1905. American Telephone and Telegraph Company, Engineering Department, 1905–19; Department of Development and Research, 1919–33. Bell Telephone Laboratories, 1933–48. Mr. Shaw's active telephone career was mainly concerned with loading problems in telephone circuits, including the transmission and economic features of the loading apparatus. The article now being published was started shortly before his retirement in 1948.