5948

THE BELL SYSTEM TECHNICAL JOURNAL

VOLUME XLV JULY-AUGUST 1966 NUMBER 6

B.S.T.J. Recommendation for Adopting the International System (SI) of Units

In 1960, the Eleventh General Conference on Weights and Measures proposed an International System of Units (designated SI System from Le Système International d'Unites) to be used when referring to physical quantities and concepts. This system was given official status in Resolution No. 12 of the General Conference when it was adopted by the 36 treaty nation members, including the United States. The SI System was adopted by the International Committee on Weights and Measures (the executive body of the General Conference) in 1962. Since then, leading professional societies, such as the American Institute of Physics (AIP), the Institute of Electrical and Electronic Engineers (IEEE), the American Society for Testing and Materials (ASTM), and other national and international organizations have adopted and are actively promoting the SI System of Units.

The Bell System Technical Journal Editorial Committee feels that a logical step toward eventual acceptance of the SI System is to use it in scientific publications wherever practical. In areas where such use is already established, no interpretative hardship should result. In applications where the use of the SI units is recent or new, it is recommended that the SI units be shown, followed by the equivalent conventional English units in parenthesis. The Committee therefore recommends that authors use the SI System alone, or with equivalent conventional English units when preparing manuscripts for publication in the Bell System Technical Journal. This recommendation is not a prerequisite to publication in the Journal; however, the Committee expects authors will adopt the SI System unless there are opposing considerations which the author feels are justifiable.

This recommendation is intended to apply only to scientific and tech-

April 12, 1967

nical data and should not be interpreted as affecting the choice of units for describing manufacturing standards or processes: i.e., inch sizes of nuts and bolts, American Wire Gauge sizes of conductors, spacing of panel mountings in inches, etc.

Information concerning the SI System is available from a number of sources (see, for example, "IEEE Recommended Practice for Units in Published Scientific and Technical Work", IEEE Spectrum, March, 1966). A list of SI units and conversion factors can be obtained by writing or calling the B.S.T.J. office.

Editorial Committee
Bell System Technical Journal

The N3 Carrier System: Objectives and Transmission Features

By G. W. BLEISCH and C. W. IRBY

(Manuscript received March 11, 1966)

The N3 carrier system development completes the second phase of a comprehensive design program to provide a new family of short-haul carrier facilities. The system includes a 24-channel, single-sideband, amplitude modulated multiplex terminal, a common carrier supply, and an N-repeatered line. Associated development effort was directed toward the provision of a shop-wired, double-bay framework to combine the carrier terminals, voice-frequency patching jacks, signaling equipment, and automatic trunk processing facilities in one equipment package.

Design objectives were established to meet today's and future stringent transmission performance requirements for direct, toll-connecting, and intertoll trunks. Taking advantage of the rapid growth in solid-state technology, significant transmission performance improvements have been achieved over earlier vacuum tube systems. In addition, installation, operating, and maintenance procedures have all been simplified. The system can be economically applied for distances as short as 35 miles, and with satisfactory transmission performance for distances exceeding 200 miles. A feature of special note is the provision of frequency correction units within the carrier terminal which essentially eliminates the frequency shift error introduced by the N-repeatered line.

I. INTRODUCTION

The Bell System has been engaged continuously in a vigorous program of carrier telephone system development since the now obsolete Type A system was introduced in 1918. This development effort has been aimed at increasing the utilization of available bandwidth, extending the operational distance, improving transmission performance, reducing transmission facility costs (initial and operating), and investigating new methods of modulation.

Prior to the end of World War II, the expense of carrier terminals and repeaters limited their application to the longer toll trunks. The type N1 carrier system,[1] which was introduced in 1950, had the design objective of providing economical telephone trunks in the 15- to 200-mile range. The N1 carrier system derived 12 voice channels and utilized double-sideband transmission over two repeatered cable pairs. A second short-haul carrier system, the type O, was developed subsequently for open wire lines. This system derived a maximum of 16 single-sideband voice channels in four groups of four channels over a single open wire pair. Terminals of the type O carrier system and the N1 carrier line facilities were then combined to form the ON carrier system.[2] The ON2 system derived 24 single-sideband carrier channels using the same line frequency space utilized by the 12 channel N1 carrier system.

In the early 1960's, a broad program was begun to redesign these short-haul carrier systems. A comprehensive description of this program has been published,[3] so an outline of the pertinent aspects will suffice here. Several factors prompted this redesign effort. The advent of Direct Distance Dialing (DDD) demanded a higher grade of transmission performance, which has helped the growth of new services such as *Data-Phone*® service. New devices and components, notably the transistor, improved ferrite inductors and transformers, and solid tantalum capacitors, offered smaller size and less power dissipation. Continued advances in circuit and system design techniques, new component mounting and equipment packaging methods, and modern manufacturing techniques gave promise that needed transmission improvements could be accomplished at reasonable cost. Indeed, the present state of the carrier development art allows the use of complex circuit designs that give significant transmission performance improvement without space or economic penalty.

The first phase of this over-all improvement program produced the N2 carrier system. A modern solid-state 12-channel double-sideband, amplitude modulated multiplex terminal was designed for use with N-repeatered lines to replace the N1 terminal. The next phase of the redesign effort was the development and production of the N3 carrier system. Here a modern solid-state 24-channel single-sideband terminal was designed for use with N-repeatered lines. The last phase of this improvement program involves the development of the N2-repeatered line. Now going into production, the N2-repeatered line will provide a level of transmission performance commensurate with that of the N2 and N3 carrier terminals.

The system aspects of the N3 carrier terminal, a compandored 24-channel single-sideband frequency division multiplex equipment, are discussed here. The paper includes a description of the system organization, a summary of transmission objectives and a review of transmission performance. In addition, certain performance parameters within the terminal are given to serve as a foundation for terminal compatibility with other manufacturers. Companion papers cover the circuit designs[4] and equipment arrangements.[5]

II. GENERAL CONSIDERATIONS

The basic function of the N3 carrier terminal is to provide a 24-channel, single-sideband frequency division multiplex for N-repeatered lines. For analog transmission, single-sideband carrier terminals normally have an economic advantage over double-sideband terminals for distances greater than 35–50 miles. Below these distances, the lower costs of the simpler double-sideband circuits and components outweigh the higher repeatered line costs per channel.

2.1 *System Transmission Performance*

The design objectives for the N3 carrier system were established to meet not only the present transmission requirements for direct, toll-connecting, and intertoll trunks but also the anticipated requirements of the future. These objectives require that the N3 terminals provide a significant improvement in transmission performance as compared to the first short-haul carrier systems (N1, O, ON). The following improvements are of particular note: wider channel bandwidth; reduced crosstalk interference, especially at low voice frequencies; better net loss stability with temperature and supply voltage variations as well as component aging; improved compandor tracking; and greater load capacity with respect to both channel signal level and system activity.

2.2 *Repeatered Line*

Planning at the very beginning of the N3 development included use of the existing N-repeatered lines of either vacuum tube or transistorized design. Coordination with existing N1, N2, and ON2 systems was required to the limited extent that all could operate within the same cable sheath. This basically required matching of carrier frequencies and carrier transmission levels with the ON2 line signal. N3-ON2 terminal-to-terminal compatibility was not deemed essential and was not provided. This approach is justified by the high growth rate anticipated for N3 channels and the significant transmission performance improvements gained in departing from terminal compatibility.

2.3 *Equipment Arrangements*

It may appear that planned use of existing N-repeatered lines would reduce the N3 carrier development activity to effort on the carrier terminal alone. In reality, two additional phases of development were carried out simultaneously with the carrier terminal work: (*i*) the provision of a common carrier supply to provide modulating and demodulating carrier frequencies and (*ii*) the provision of an integrated equipment arrangement including carrier terminals, in-band signaling units, 4-wire voice-frequency (VF) patching facilities, and automatic trunk processing equipment.

Costs were continually reviewed throughout the development. As expected, the desired transmission improvements generally increased equipment costs. However, the packaging of the carrier terminal, 4-wire VF patching facilities, in-band signaling unit mounting arrangements, and trunk processing equipment *into an integrated shop-wired bay frame*, made it possible to more than offset the increased equipment costs by savings in floor space, engineering and installation. In addition, the solid-state circuits offer noteworthy power savings and should make possible considerable savings in maintenance expense.

2.4 *Signaling*

The modern in-band single frequency signaling family makes available the wide range of signaling options needed to provide the requisite flexibility in trunk supervision. This signaling system makes N3 channels compatible with channels or trunks provided by long-haul carrier systems. Signaling economies are achieved from this whenever a trunk is built up of two or more channels in tandem, since signaling terminals are required only at the trunk ends, and none are required at the channel junctions. Also, it was appreciated that the use of in-band signaling would ease circuit design limitations associated with providing increased channel bandwidth; the 3700-Hz out-of-band signaling tones as used in N1, O, and ON systems are necessarily transmitted at high level to insure adequate signal-to-noise performance. Since a high loss in the voice circuits is required to discriminate against the signaling frequency energy, this seriously limits the voice-frequency bandwidth that can be achieved.

2.5 *Frequency Precision*

The low frequency channel response objective and associated discrimination requirements for the channel filters dictated that the frequency shift of the sideband signal with respect to the channel filter be limited to 20 Hz. This order of precision was required since voice-frequency equalization of the channel filter band-edge roll-off was desired to improve adjacent channel suppression. If effective voice-frequency equalization was to be achieved, the received band had to remain within the channel filter frequency allocation within the 20-Hz tolerance. This made clear the need to eliminate the frequency shift introduced by the N-repeatered line, which can be as large as 100 Hz. In order to achieve the desired transmission improvements, it was necessary to incorporate frequency correction units in the terminal design which essentially eliminate this source of frequency error.

Further studies, assuming no repeatered line frequency shift, indicated that the long term frequency stability attainable with economically practical independent oscillators for the terminal modulators (similar to those used in existing Bell System short-haul carrier systems) was marginal for single-sideband operation, particularly with respect to temperature variations. Economic considerations led to the provision of a common carrier supply rather than stabilized individual oscillators. The primary frequency source stability specified was ± 7 parts per million over a six month period. While this value is substantially better than the precision required for message trunk considerations, it was deemed necessary for possible future program channels and other special service applications.

III. SYSTEM ORGANIZATION

3.1 *Modulation Plan*

The modulation plan adopted for the N3 system derives the 24-channel line signal within the terminal from two 12-channel groups in two steps of modulation. Fig. 1 illustrates the frequency allocation. Twelve channel filter designs, each with a 4-kHz upper sideband allocation, are provided in the frequency range 148–196 kHz. Each set of twelve channels forms a channel group. Each channel group is then group-modulated (280 or 232 kHz) to form the N-carrier low group

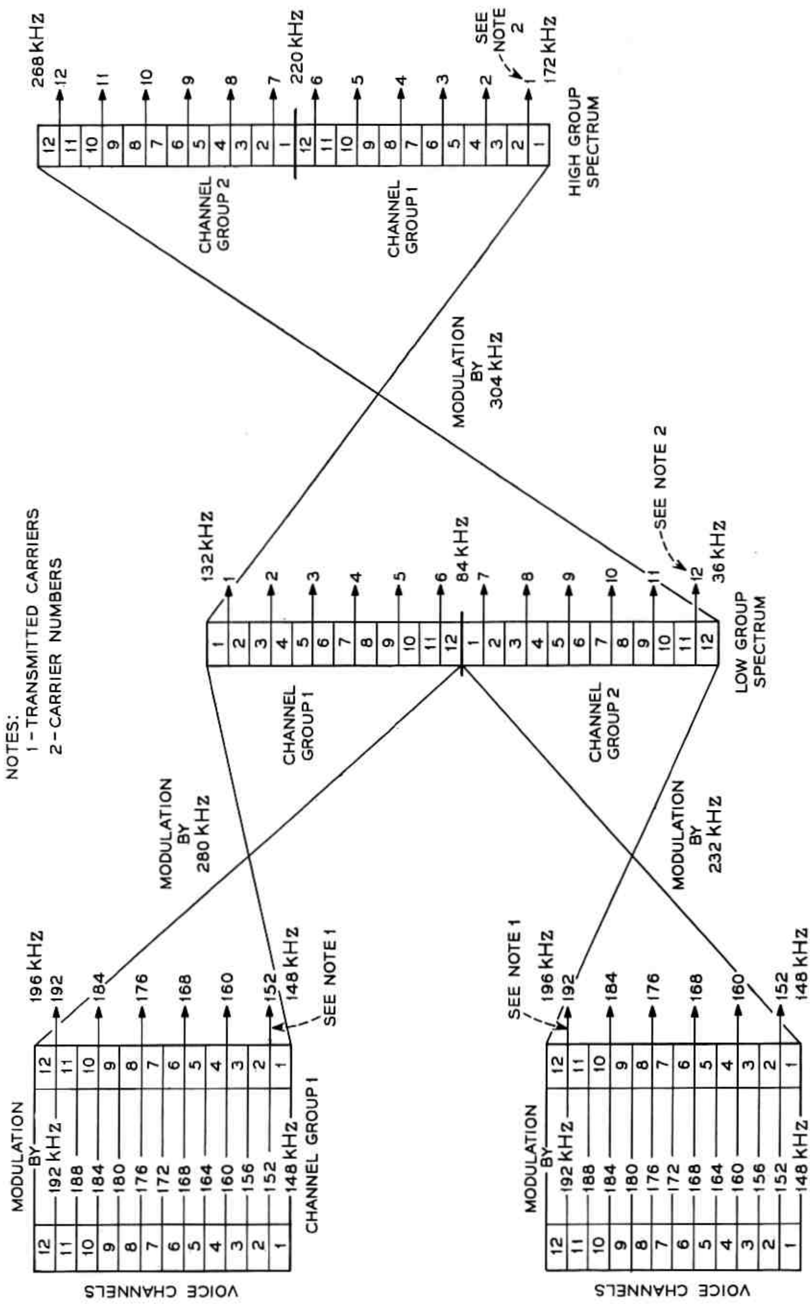

Fig. 1 — N3 carrier system frequency allocation.

band (36–132 kHz). Depending on the office location, either the 24-channel low group signal is fed directly to the line or the signal is given an additional step of modulation (304 kHz) to form the N-carrier high group band (172–268 kHz). The carriers associated with even channels are transmitted on the line to accommodate total power regulation of the type N repeaters; they are inserted separately for each channel group in the 148–196 kHz band.

The two 12-channel group organization of the N3 terminal was guided by three major factors: (*i*) the six, 4-channel group organization of the ON2 system, dictated by the use of O-carrier channelizing equipment, was uneconomical in terms of the excessive number of group modulators and amplifiers required; (*ii*) present rates of circuit growth normally can justify additions in units of twelve rather than four channels; (*iii*) field requests had indicated the desirability of interconnecting short and long-haul carrier facilities at group rather than voice frequencies. Sideband orientation was chosen on the basis of (*iii*) above. In the 148–196 kHz band, channel filters select the upper sideband. A single step of group modulation with a carrier at 256 kHz could translate the N3 channel group band to the 60–108 kHz lower sideband oriented group signal long established in the long-haul plant. Such group interconnecting equipment for short and long-haul systems is now being developed.

3.2 *Terminal Description*

The N3 carrier system, as illustrated in Fig. 2, is comprised of a transmitting and receiving terminal, carrier supplies, and an N-repeatered line. While the over-all N3 development included the packaged double bay arrangement, the VF jack field, signaling, and trunk processing components are not normally considered part of the carrier system.

For message service the signal in each of the 24 channels in the two 12-channel groups is compressed at a syllabic rate and modulated into the 148–196 kHz band. Each of the twelve channel filters passes the upper sideband, provides some suppression to the carrier, and rejects the lower sideband and other products of the modulator output. A balanced resistive multiple combines the twelve single-sideband filter outputs. The transmitted carriers (for the even channels) are also inserted at this multiple. Each 12-channel group is then modulated into its portion of the low group N carrier band.

A hybrid circuit combines the two channel group signals providing a continuous 24-channel spectrum with inserted transmitted carriers in the low group N-frequency range of 36–132 kHz. The group transmitting unit slope equalizes and amplifies the signal for line transmission. Both low and high group transmitting units are available; the high group unit contains an additional modulator to provide the frequency translation.

A line terminating unit is placed between the group equipment and the line pairs. The line terminating unit includes provisions for: (*i*) powering the repeaters in the line power section (up to three repeaters) adjacent to the terminal, (*ii*) the insertion of optional input and output span pads to adjust receive and transmit levels, and (*iii*) terminal secondary lightning protection for both line pairs.

In the receiving terminal the modulation steps are reversed. The group receiving unit slope equalizes, amplifies, and regulates the signal received from the line. Both high and low group receiving units are available; the high group receiver contains an additional modulator for frequency translation. At the output of either group receiving unit, the composite 24-channel signal is in the low group N-frequency band. The output signal is fed to two channel group demodulators via a balanced resistive splitting network.

Each channel group demodulator translates the appropriate portion of the low group N-carrier band to the 148–196 kHz band. The output of each channel group demodulator is fed to six double-channel regulators; in addition, the terminal frequency correction and alarm units are bridged at this point.

Frequency correction units are provided separately for each channel group. The function of the frequency correction unit is to derive the channel group demodulator carrier frequency. In order to eliminate the frequency shift introduced by the N-repeatered line, this derived channel group demodulator frequency is offset from its nominal value by an amount equal to the line frequency shift.

Separate alarm units are provided for each 12-channel group; these function not only to determine a carrier failure but also to determine when carrier transmission is satisfactory for service restoral.

The double-channel regulators automatically adjust the level of the two channels immediately adjacent to a transmitted carrier. A double-channel regulator is associated with each transmitted carrier since amplitude distortion in the repeatered line can create a substantial difference in the level of the received carriers. A channel demodulating carrier signal for the associated even channel is also obtained from this regulator. Odd channel demodulating carriers are obtained from

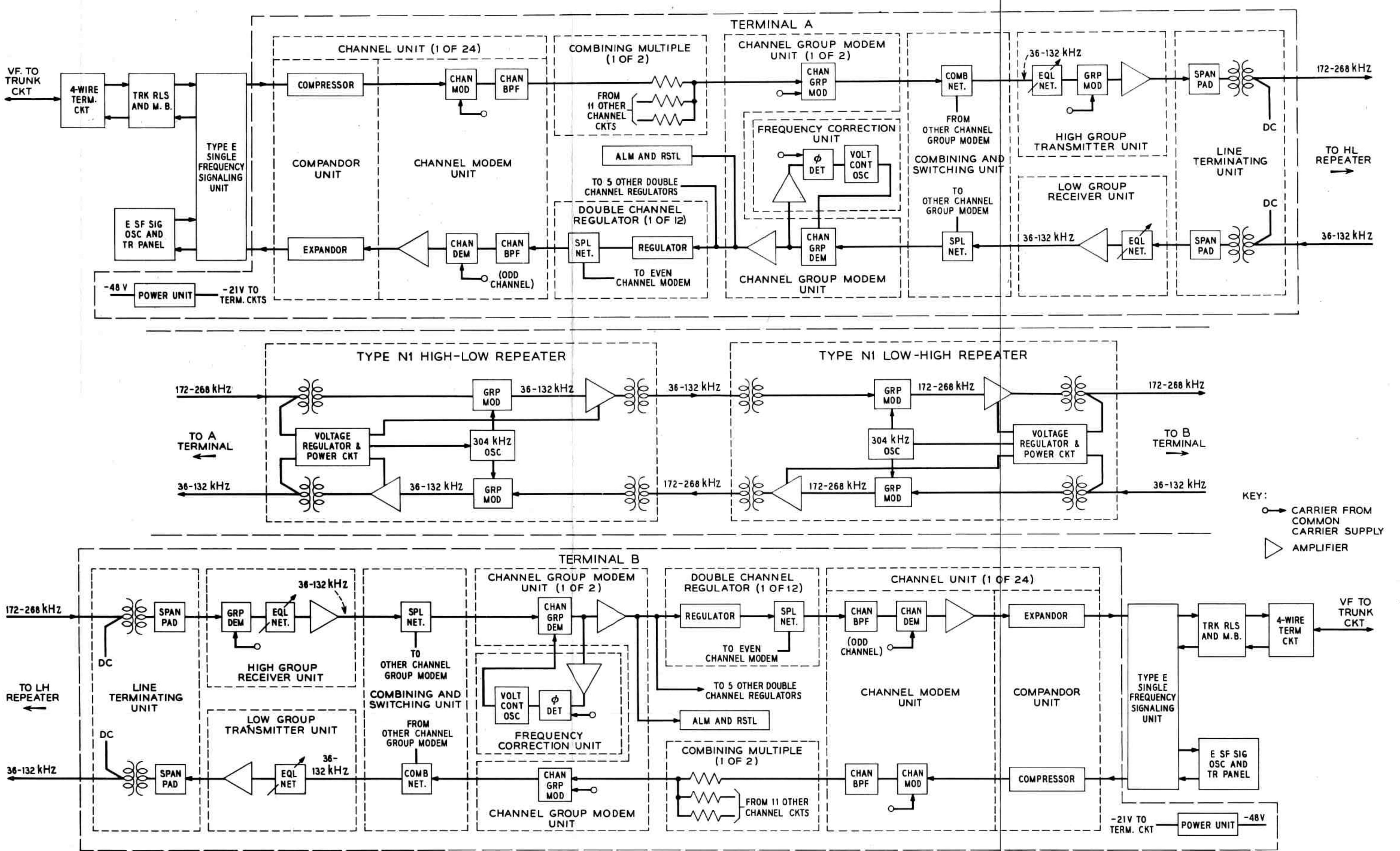

Fig. 2 — N3 Carrier system block diagram.

the common carrier supply. Each double-channel regulator signal output is fed to two channel demodulators.

The channel demodulator circuits include the receiving channel filter, demodulator, amplifier, and low-pass filter. The amplifier includes a feedback equalizer to compensate for band-edge amplitude distortion of both transmitting and receiving channel filters; the low-pass filter provides a peak of suppression at 4 kHz to suppress any tone resulting from the transmitted carriers. To complete the N3 channel, an expandor restores the original level range.

3.3 *N-Repeatered Line*

Essential features* of the N-repeatered line are also illustrated in Fig. 2. Originally developed for the N1 carrier system, the N-repeatered line is now also used for Western Electric ON1, ON2, N2, and N3 carrier systems. A separate cable pair is employed for each direction of transmission, usually within the same cable sheath. Further electrical separation of these signals is obtained by using different frequency bands for each direction of transmission; 36–140 kHz (low group) for one direction and 164–268 kHz (high group) for the other direction. These high and low group bands allow frequency space for 26 single-sideband (13 double-sideband) channels. In the N3 carrier system application, however, only 24 channels are transmitted within the fixed frequency bands 36–132 kHz and 172–268 kHz.

Frequency frogging, a feature of the N-repeatered line, involves the interchange of high and low group frequency bands at each repeater. Two types of repeater equipment units are employed; the first is called high-low (HL) and the second low-high (LH). Both types of repeaters interchange the group bands, each including a modulator for each direction of transmission and a common local oscillator operating at 304 kHz. Frequency frogging blocks a major circulating cross-talk path around each repeater and provides first-order equalization of line slope (increasing attenuation at higher frequencies). Other methods of slope control are also available including fixed slope networks for significant slope equalization and a slope switch on each repeater for small slope adjustments. In combination, these slope controls allow the transmission engineer to minimize line noise by design.

Primary power for N-type repeaters is either provided locally or transmitted over the cable pairs. For the transistorized designs, as many

---

* More detailed descriptions and objectives of N-repeatered lines can be found in Refs. 1 and 3.

as three repeaters in tandem can be powered via simplex circuits on the cable transmission pairs.

3.4 *System Levels*

Certain restraints on the transmission levels of the N3 carrier system were imposed by the nature of the development. As previously discussed, use of N-repeatered lines of existing design and the need for coordination with other N and ON systems operating within the same cable sheath clearly defines the required levels on the line side of the terminal. The compandor design chosen is the same as that used in the N2 carrier system; this defines the levels throughout the compandor circuit.

Fig. 3 indicates the transmission levels at major functional points throughout the N3 system. The voice-frequency and sideband values shown are those which would be *measured* if a 0 dBm, 1-kHz sine wave tone were applied at the zero transmission level point (0 TLP). This means that the levels shown between the output of the compressor variolosser and the input of the expandor variolosser are the actual, compressed levels.*

For that portion of the carrier system where the transmitted carriers are present, (including all of the repeatered line) system levels are quoted on the basis of the transmitted carrier amplitude.

The channel sideband power transmitted on the line (compressed) for a 0 dBm test tone at the 0 TLP is nominally 3.5 dB below the transmitted carrier power. This is the same carrier-to-sideband ratio employed in the ON2 system. (Historically, this carrier-to-sideband ratio was first employed in type O systems to minimize static noise.) The single sideband has 9 dB more power than is contained in each sideband of the double-sideband N carrier systems. The two sidebands of the double-sideband systems add in phase while the noise powers, being noncoherent, add on a power basis; this results in a 3-dB signal-to-noise advantage for a double-sideband system over a single-sideband system with equal sideband amplitudes. Since the N3 single-sideband power is 9 dB greater, a theoretical 6-dB signal-to-noise advantage is obtained. This has proved advantageous since the single-sideband systems are usually applied on the longer circuits.

Single-sideband systems such as N3 do not have the overmodula-

---

* In Ref. 1, fictitious "message level" values were given as an indication of what the level would be if the 0 dBm test signal at the 0 TLP was not compressed. Such values can facilitate computations but have been omitted from Fig. 3 for simplicity.

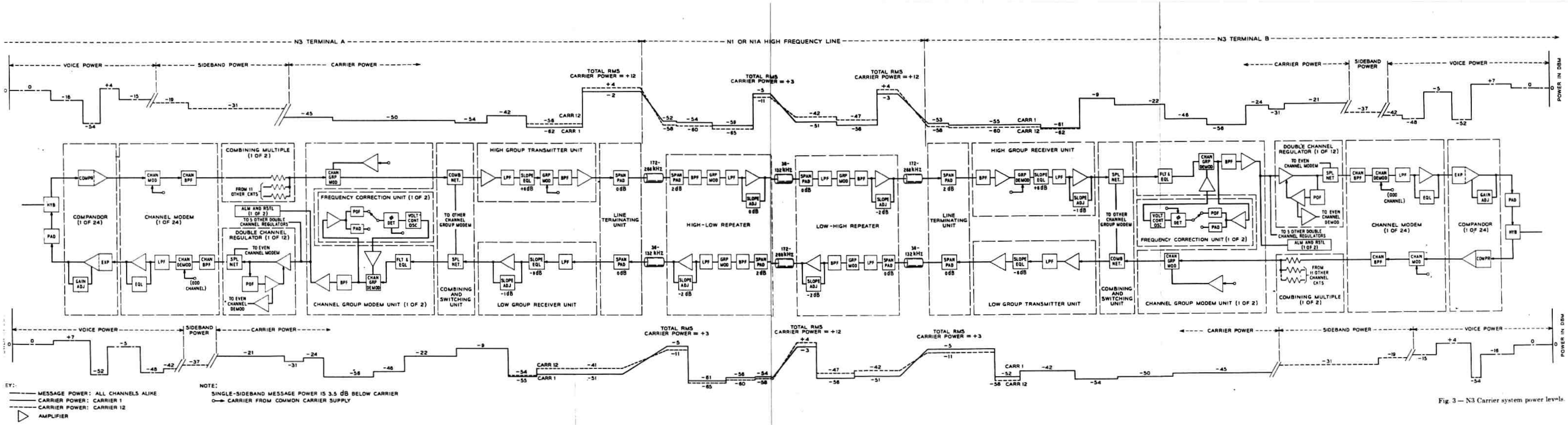

Fig. 3 — N3 Carrier system power levels.

tion restrictions inherent with double-sideband modulation. Thus, the carrier-to-sideband ratio employed does not of itself limit the load capacity of the channel. In the case of O, ON, and N3 systems, the carrier-to-sideband ratio has been chosen to maintain the line signal power reasonably constant with expected variation in system activity. Because of the total power regulation used in N-type repeaters, variations in activity result in second-order changes in the signal-to-noise performance.

Where design choice existed, signal levels within the N3 terminal were chosen as a compromise between a high level to keep the signal above noise and a low level to avoid nonlinear distortion. With normal system and channel loading the controlling noise source of the system is crosstalk and impulse noise on the N-repeatered line. For back-to-back terminals, the controlling noise source is the first circuit noise of the expandor amplifier on compandored channels and modulator noise in the high group transmit or high group receive units for noncompandored channels; some external noise from the power supply is also measurable on noncompandored channels.

One of the sources of noise is the modulation performance of the group equipment. All of the modulating and transmitted carriers of the N3 system are developed from a common carrier supply.[4] Therefore, all transmitted carriers have a fixed relationship with each other and the peak amplitude of the carrier tones can be periodically large. In addition, the peak amplitude of the sideband powers can also be large if several channels are transmitting the same signal. The sum of these peak amplitudes imposes a severe load handling requirement on each of the group units of the N3 terminal. To reduce this power handling requirement, the phase of certain transmitted and channel modulating carriers have been reversed with respect to others. Better modulation performance of the group equipment results.

Nonlinear distortion on compandored message channels is controlled primarily by the variable loss "variolosser" elements in the compandor. For noncompandored channels, the channel demodulator amplifier contributes the most nonlinear distortion.

IV. SYSTEM FEATURES

4.1 *Compandor*

Short-haul carrier systems have enjoyed rapid growth in the Bell System since the introduction of type N1 carrier in 1950. At the end of 1965, over 700,000 two-way, short-haul channels derived by frequency division multiplex had been placed in commercial service. This is now nearly two-thirds of all frequency division carrier channels in the Bell System. A major factor in this widespread acceptance has been the relative ease in adapting a great mass of cable pairs, most of them originally installed for VF transmission, for short-haul carrier use. A key factor in this ease of application is the built-in compandor, a feature of all previous short-haul carrier systems retained in type N3. By compressing the volume range transmitted and by making a corresponding expansion of the received volume range, a compandor affords a substantial increase in the amount of crosstalk and noise which can be tolerated in the carrier frequency part of the system. This includes the entire line transmission facility as well as most of the carrier terminal. A compandor avoids the need for expensive line treatment such as crosstalk balancing or short repeater spacing and effectively enhances band filter discrimination.

The compandor design chosen for N3 is the same as that employed in the N2 carrier system.[7] This choice was made with the benefit of some operating experience of the N2 compandor in the field; excellent stability, harmonic distortion, and tracking performance observed on early N2 systems gave promise (which has since been verified) that needed improvements in these areas over early short-haul carrier systems could be achieved with the chosen design.

4.2 *Channel Filters*

The over-all performance of a single-sideband, amplitude modulated, frequency division multiplex carrier terminal is influenced in large measure by the channel filter discrimination and in-band amplitude distortion characteristics. The two-section, quartz crystal channel filters of new design are a major factor in the improved performance of the N3 carrier terminal. The insertion loss characteristic of a typical channel filter is illustrated in Fig. 4.

As opposed to the twin channel (one upper, one lower sideband about a common carrier) arrangement used in previous O and ON carrier system designs, the N3 carrier terminal employs the same sideband orientation for all channels. This was chosen partly to achieve a substantial improvement in adjacent channel crosstalk, particularly at low voice frequencies. In the twin channel arrangement, the high-energy, low frequency portions of the speech spectrum of both channels are clustered closely around the common carrier. Any vestige of the

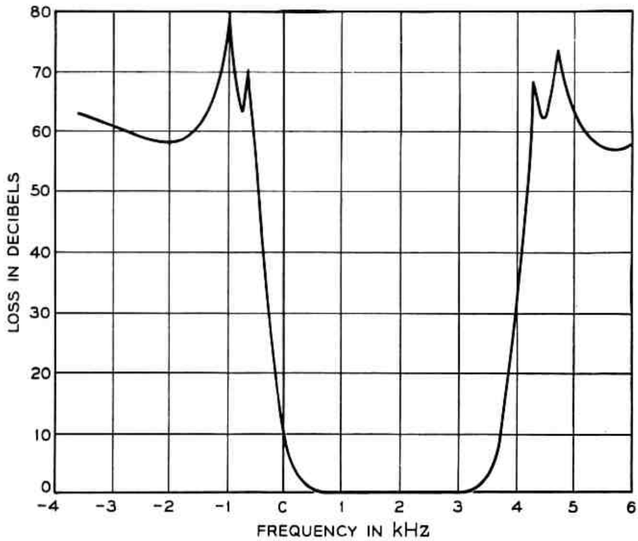

Fig. 4 — N3 channel filter discrimination.

unwanted sideband falls upright in the adjacent channel, placing most stringent requirements on the channel filter discrimination shape in the narrow frequency range between pass and stop bands. Common orientation of sidebands separates the high energy portions of adjacent channels to the full extent of the channel spacing; this provides about 10 dB less crosstalk interference than the twin channel arrangement, assuming speech type signals and the same discrimination shapes for both. In addition, the interference for the common sideband orientation arrangement is inverted in the disturbed channel, even further reducing the effect of this form of crosstalk.

The decision to depart from the earlier twin-channel arrangement ruled out the possibility of terminal-to-terminal compatibility with ON2. Yet the excellent crosstalk performance reported in Section 6.5 for back-to-back terminals makes clear the sacrifice of terminal compatibility was not without reward.

4.3 *Net Loss Stability*

Particular care was taken throughout the N3 terminal design to assure a high order of net loss stability. Liberal use of negative feedback, close control of compandor tracking, the use of a regulated power converter, group and double channel regulation, precise regulation of the transmitted carrier insertion level and temperature compensation of networks all contribute to the excellent stability attained.

No transmit level adjustment is provided and some variation in the carrier-to-sideband ratio is tolerated from channel-to-channel as a result of manufacturing tolerances. The absolute power of each transmitted carrier is factory adjusted and maintained at the regulator output to within ±0.05 dB of a nominal value. Variation in the carrier-to-sideband ratio is acceptable within limits as long as it is not time variant. Temperature equalization is provided to compensate for the transmitting crystal channel filter loss variations which are in the order of 0.015 dB per degree Fahrenheit.

Having established the carrier-to-sideband ratio at an early point in the transmitting terminal, one can expect the same ratio to be maintained throughout the broadband terminal circuitry and repeatered line. From a general stability point of view, the major consideration involves maintaining the frequency spectrum essentially flat for each transmitted carrier and its two associated sidebands (a bandwidth of 8 kHz). This was accomplished by incorporating amplitude equalizers in several of the group and channel group filters, reducing the ripple distortion in these units to less than 0.1 dB over any 4-kHz increment of the band.

The N-repeatered line is a factor in both the stability and channel frequency response performance of the system. Any amplitude distortion which exists over the bandwidth of a channel is reflected in the channel frequency response; any time variant change in the amplitude distortion which exists over the two sidebands associated with each transmitted carrier results in correlated level changes within the channels. A well-engineered and maintained N-repeatered line is essential to obtain the superior transmission performance built into the N3 terminals.

The double-channel regulator automatically maintains the received carrier level essentially constant at its output. Since the carrier-to-sideband ratio is fixed, the associated channel sidebands are also regulated. This accommodates flat loss changes in the received signal. Wide range and extreme regulation stiffness are provided. Part of the need for good regulation stems from the use of common sideband orientation; in the odd channels, a 1-kHz test tone is separated from the transmitted carrier controlling its regulation by 3 kHz. This emphasizes the need to reduce the amplitude distortion of the repeatered line and terminal group equipments outlined in the above paragraph. Each double-channel regulator also supplies the channel demodulator carrier signal for its associated even channel. Two elements within the double-channel regulator are subject to significant temperature variation: (*i*) a voltage reference diode and (*ii*) the narrowband crystal pick-off filter. These elements have opposite effects and a net temperature compensation is accomplished with a single temperature compensating resistor.

In the receiving channel demodulator circuit temperature compensation is also provided to compensate for the receiving crystal channel filter variation.

The expandor in the receiving terminal contains the only operating level adjustment provided for each channel. Its purpose is to accommodate the small level variations which accrue from manufacturing tolerances and length differences in central office cabling as well as the inherent long term aging. Once set there is little likelihood that it will need readjustment for at least six months.

4.4 *Frequency Correction*

As a result of the frequency-frogging process in each N repeater, small errors in the 304-kHz repeater oscillators tend to shift the transmitted line signal from nominal by the amount of frequency error. These small errors at each repeater can accumulate, and line frequency shifts approaching 100 Hz have been observed on long repeatered lines. Frequency correction units are employed in the N3 terminal to essentially eliminate this repeatered line frequency shift.

The line frequency shift is eliminated* with the aid of a phase-locked-loop which compares a received carrier signal at the output of the channel group demodulator with the correct frequency generated in the office primary frequency supply. Departure from phase coherence of these two signals is used to control a variable frequency oscillator within the frequency correction unit. This oscillator provides the correct channel group demodulator carrier frequency which will synchronize the selected received carrier to the carrier frequency supply at the receiving terminal.

* Early N3 frequency correction units eliminated the line frequency shift by selecting a particular carrier by means of a pick-off filter and modulating this signal with a precise local carrier. The upper sideband of this modulation process contained the nominal channel group demodulator frequency plus the line frequency shift. Using this carrier signal for channel group demodulation resulted in a desired lower sideband output without frequency shift. The transmitted carrier pick-off filter requirements are rather severe, the range of frequency shift accommodated being compromised with desirable crosstalk objectives. The new frequency correction unit, based on the phase-lock principle, has eliminated the need for this compromise.

4.4.1 *Channel Frequency Stability*

The elimination of the line frequency shift makes it possible to obtain odd channel demodulating carriers from the carrier frequency supply at the receiving terminal. The only frequency error present in the odd channels after correction is due to the independence of the carrier frequency supplies at the transmitting and receiving terminals; this is controlled by specifying a long term (six month) stability and maintenance limit of ± 7 parts per million for each primary supply. Since the odd channels of the N3 terminal are demodulated with locally generated carriers, frequency differences between the primary frequency supplies at transmitting and receiving terminals result in a frequency error in the detected signal. With the two primary frequency supplies at opposite tolerance extremes, the maximum frequency error in an odd channel, demodulated voice-frequency signal is 0.6 Hz. Even channels are coherently detected with their associated transmitted carrier and no frequency error is present.

4.5 *Voice-Frequency Equalization*

The precise control of frequency within the N3 terminal allows voice-frequency equalization of the carrier frequency amplitude distortion caused by the channel filter roll-off at the band edges. Equalization is provided at both band edges for amplitude distortion introduced by the transmit and receive channel filters with a network in the receiving channel demodulator amplifier.

4.6 *Alarm Provisions*

The failure of an N-repeatered line or an N3 terminal group unit causes the loss of 24 carrier derived trunks; failure in a 12-channel group equipment or a frequency correction unit causes the loss of 12 trunks. Rather extensive carrier alarm features have been incorporated in the N3 system design to minimize the effects of such failures on customer service and office switching equipment. Separate carrier alarm units are provided for each twelve channel group; these function independently, not only to determine a carrier failure, but also to determine when transmission is satisfactory for service restoral. The terminal alarm units control auxiliary trunk release and make-busy panels which "condition" the trunks during the failed interval by dc control of the trunk supervision leads. Trunk conditioning during a carrier failure interval first involves making the trunk idle to stop subscriber charges and, for most trunks, automatically disconnecting the subscriber. Subsequently, the trunks are made busy to avoid selection of an unusable trunk by the switching machine. The trunk is held in the busy state until transmission is satisfactory for service restoral; at such time the trunk conditioning is removed thus making it available for service.

The carrier alarm units are bridged on the output of the channel group demodulator circuit and register an alarm if the received power drops below threshold for more than about two seconds. Following registration of the alarm, a forced transmission failure in the same channel group is induced toward the far* terminal and a control signal to condition the associated trunks to the idle state is maintained for about 10 seconds. Note that the forced transmission failure results in the registration of an alarm in the same channel group at the far terminal. At the end of this 10 second interval a control signal to condition trunks to the busy state is generated and maintained for the duration of the carrier failure; also the automatic restoral sequence is initiated.

The automatic restoral sequence is controlled by the alarm unit as it applies and monitors transmission test tones on two of the channels (which are out-of-service during the failure interval). Recalling that alarms are registered at both near and far terminals, a test tone is applied on the first test channel of the near terminal which is monitored at the far terminal; also, a test tone applied on the first test channel of the far terminal is monitored at the near terminal. Within both alarm units the signal-to-noise ratios are evaluated. When the signal-to-noise performance in one direction is satisfactory for service restoral, and such performance is maintained for a period of about 10 seconds, tone is applied on a second test channel. In the case of a unidirectional transmission failure, the signal-to-noise performance is satisfactory in one direction and unsatisfactory in the other. Service restoral is withheld in this instance since the terminal receiving satisfactory transmission does not have an indication on the second test channel that transmission is satisfactory in the other direction. When the fault is cleared, both terminals have an indication that transmission has been satisfactory in both directions. At essentially that instant, both terminals are restored to service by removing the transmission test tones and trunk conditioning.

In connection with the automatic trunk conditioning provisions, an

* Alarm arrangements at the two terminals of an N3 carrier system are identical. Because of this symmetry, the terms *near* and *far* are used for terminal distinction

option is available to allow manual overriding of the trunk conditioning. Such overriding of conditioning permits trunks with this applied option to be restored to service during the carrier failure interval by patching to alternate transmission facilities.

4.7 *Special Service Provisions*

In addition to the conventional use of compandored N3 channels for direct, toll-connecting, and intertoll message trunks, most of the channels are satisfactory for Schedule C & D Program service. Exceptions* are the end channels 1 and 2 of channel group one and end channels 11 and 12 of channel group two. End channel restrictions are imposed as a result of channel frequency response variations due primarily to repeatered line characteristics.

For noncompandored channel applications a VF amplifier is available to replace the compandor. One VF amplifier application includes its use at the junction of two N3 channels wired in tandem to derive one over-all compandored circuit; use of the VF amplifiers avoids the use of two compandors in tandem.

Another common use is in the provision of private line, voice band data transmission circuits; for such signals, the signal-to-noise improvement obtained with a compandor is limited and a more economical circuit is obtained with the VF amplifier.

4.8 *Maintenance and Testing*

Rapid growth in the number of carrier channels and today's stringent transmission performance requirements are but two of many factors emphasizing the importance and need for simplified maintenance of a modern carrier system. In the N3 carrier development it was possible to incorporate many of the operating convenience and maintenance innovations originally introduced and field proven with the N2 carrier system. Included are the automatic carrier alarm and trunk processing features, the use of a regulated power supply, separate line terminating units incorporating repeater power feed circuits and plug-in, flat loss span pads, provision of easily replaceable slope equalizing

* On systems utilizing the original frequency correction unit design, channel 6 of channel group 1 and channel 2 of channel group two were also excepted for Schedule C & D Program service. These are the channels associated with the frequency correction carriers; the restrictions are necessary to avoid crosstalk resulting from the high-energy, low frequency components of program material from entering other message channels.

networks in the group transmitting and receiving equipment, and the use of a single gain control for the adjustment of channel net loss.

The N3 terminal requires 15 distinct carrier frequencies obtained from a common carrier supply which can accommodate as many as 26 N3 terminals totalling 624 channels. The carrier supply arrangement allows the optional provision of alternate spares with automatic alarm and switching features to assure continuity of service. Active N3 terminal units handling 24 channels are provided with in-service switching capability; a terminal switching set allows essentially hit-free* in-service switching of the group transmitting unit, group receiving unit and power supply. This switching set also contains an accurate voltmeter which allows precise adjustment of the regulated power supply voltage.

In-service signal level or voltage measurements may be made on compandors, modems, group transmitting and receiving units, and the power converter by means of pin jacks on the front of the units or by connection to switching jacks. Similar pin jacks on the group transmitting and receiving units, channel group modem unit, and frequency correction unit permit in-service transistor emitter voltage measurements. On an out-of-service basis, a portable test stand allows terminated measurements on the compandor, voice-frequency amplifier, channel modem or alarm unit, as well as providing bridging access to all input and output connections of these units.

V. OBJECTIVES

5.1 *Broad Objectives*

As previously stated, the broad objective of the N3 carrier system development was to provide a modern replacement for the ON2 carrier system with improved performance. More specifically, a single-sideband, amplitude modulated frequency division multiplex terminal for N-repeatered lines was desired. Transmission performance capable of meeting today's stringent requirements for direct, toll connecting and intertoll trunks was a must. These are essentially the same broad objectives set forth for the N2 carrier system with the exception of single-sideband modulation and the replacement of the ON2 carrier system. Hence, the transmission performance objectives for the N3

* During the switching transient for group transmitting and receiving units, a 1 dB level reduction occurs for a duration of about 7 milliseconds; the phase transient during this interval is negligible.

TABLE I — N3 CARRIER SYSTEM — TRANSMISSION PERFORMANCE OBJECTIVES

Channel gain-frequency response (−3 dB points)	200–3450 Hz
Amplitude Distortion: 600–3000 Hz	±1.0 dB
Net loss stability (six months)	
Distribution grade (standard deviation)	0.5 dB
Bias (average)	0.25 dB
Short term loss variations — "beats" (intrasystem)	0.1 dB peak-to-peak
Compandor tracking (+8 to −40 dBmO)	±1.0 dB
Compandor tracking (+10 to −52 dBmO)	±2.0 dB
Compandor advantage (average/minimum)	30/28 dB
Channel noise* (at 0 TLP)	
Idle terminal (compandored)	16 dBrnC
Idle system (compandored over 100 miles)	26 dBrnC
Loaded terminal, (OVu, 40% activity, noncompandored)	48 dBrnC
Loaded terminal, (OVu, 40% activity, compandored)	20 dBrnC
Channel distortion	
2A − B; 0 dBmO fundamentals — dB below fundamentals	30 dB
Crosstalk — all terminal sources	70 dB
Equal level coupling loss	
Repeatered line amplitude distortion	
Slope across channel bandwidth 40 repeaters, 200 miles (compressed)	±0.5 dB

* Measured values are given; effective values are usually considered to be about 5 dB greater due to syllabic operation.

carrier system, as summarized in Table I, are quite similar to those for N2.

5.2 *Detailed Functional Objectives*

The over-all performance of the N3 carrier system is controlled by both the N3 carrier terminal and the N-repeatered line. Performance objectives for the various functional blocks of the N3 terminal were determined on the basis of analytical studies. For the N-repeatered line, the N2 repeater requirements[3] were assumed. Use of present N-repeatered line facilities is permissible and it is anticipated that with few exceptions, all transmission performance objectives can be met. One significant exception is the repeatered line amplitude distortion objective which is difficult to maintain on long systems using repeaters of existing design.

Substantial analysis effort was directed toward determining specific performance objectives for the various functional blocks of the N3 carrier terminal. Some of the areas covered by this work include:

(*i*) Study of the various interference mechanisms.

(*ii*) Consideration of today's state of the art, in particular with respect to filter network, modulator and amplifier preformance.

(*iii*) Selection of a judicious allocation of the over-all performance objective among the contributing functional blocks.

Table II presents some of the specific performance objectives derived with this procedure.

VI. TRANSMISSION PERFORMANCE

The various components of a carrier system may have satisfied their individual design objectives, but the transmission performance of all of the assembled individual components is the final criteria as to how well the over-all system objectives have been met. The measured transmission performance of manufactured N3 carrier systems shows that all objectives for the system have been satisfied with ample margins.

The transmission results given here represent the performance of the

TABLE II — FUNCTIONAL UNIT PERFORMANCE OBJECTIVE

Unit	Parameter	Objective
Channel modem		
Modulator	Carrier leak(*i*)	−32 dBmO
Modulator, demodulator	Linearity(*ii*)	−54 dBmO
Receive amplifier	Linearity (*ii*)	−46 dBmO
Channel group modem	Linearity(*iii*)	−110 dB
Group transmitter	Linearity(*iii*)	−95 dB
Group receiver	Linearity(*iii*)	−100 dB
Frequency correction unit	Distortion(*iv*)	−55 dB
	Freq. shift range	±100 Hz
Alarm unit	Alarm threshold(*v*)	−18 ± 2 dB
Double channel regulator		
Through transmission amp.	Linearity(*iii*)	−105 dB
Regulation	Range/stiffness(*vi*)	±12 dB/±0.25 dB

Notes:
(*i*) excludes channel filter suppression
(*ii*) (2A − B) distortion product resulting from 0 dBmO fundamentals
(*iii*) (2A − B) distortion product relative to transmitted carrier levels.
(*iv*) second harmonic content relative to fundamental carrier output level.
(*v*) received power relative to nominal
(*vi*) output variation relative to nominal at extremes of regulation range.

N3 carrier system as measured between the VF IN and VF OUT jacks of the N3 packaged bays or their equivalent in a centralized patch bay. In addition, these results include certain interface data within the terminal. Such data will be helpful to manufacturers for non-Bell companies in developing carrier terminals compatible on an end-to-end basis with N3 terminals. Items included for this reason are compressor input-output characteristics, compressor intramodulation figures and carrier-to-sideband ratios for channel voice-frequency inputs from 200–3500 Hz.

6.1 *Frequency Characteristic*

The gain-frequency characteristics of N3 channels, illustrated in Fig. 5, are representative of present N3 carrier product on a system of about average length having a good high frequency line characteristic. In general, these N3 channel characteristics compare favorably with those of the A5 channel bank and the N2 carrier channels.[3] The principal concern has been the positive peaks in the response near the band edges of the characteristic. The return losses of associated VF circuit equipment outside of the N3 carrier system at the band edge frequencies can lead to near-singing distortion if the peaks are not controlled. Some average bandwidth at the lower frequencies of the N3 channel has been sacrificed in order to minimize these gain peaks.

The design objective for the carrier-to-sideband ratio of a 0 dBmO,

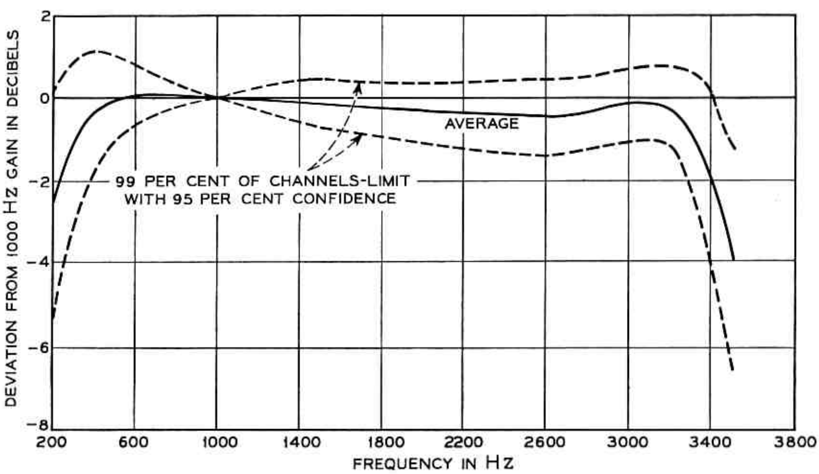

Fig. 5 — N3 channel VF gain-frequency characteristic.

1000-Hz compressed tone at the output of the N3 group transmitting unit is 3.5 dB. The carrier-to-sideband ratio for other voice frequencies that can be transmitted by the channel may vary slightly from this value, reflecting primarily the response characteristics of the channel filter of the N3 channel modulators. The mean carrier-to-sideband ratio for a 1000-Hz, 0 dBmO tone was 3.27 dB. Statistical results indicate that carrier-to-sideband ratios between 2.52 and 4.07 dB would encompass 99 per cent of the channels with 95 per cent confidence.

Fig. 6 eliminates the flat loss deviations from the carrier-to-sideband ratios and shows the transmitted sideband levels for all voice frequencies transmitted in the channel with respect to the sideband level for 1000 Hz. The characteristics shown in Fig. 6 essentially represent the passband frequency characteristic of the transmitting channel filters. The receiving channel filters are identical. The variation of the average curve of Fig. 6 for frequencies at the edge of the voice band multiplied by four (two filters and an expansion ratio of 2:1) will be worse than the deviations of the average curve for these same frequencies shown in Fig. 5. The equalizer at the receiving end of the N3 channel provides the difference between the two results at the low and high frequency ends of the voice channel.

6.2 *Compandor Tracking*

Compandor tracking performance of the N3 channels is indicated by the curves of Fig. 7.

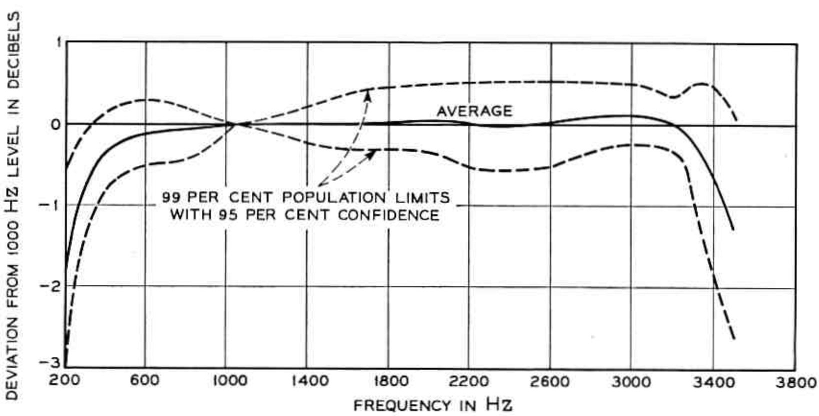

Fig. 6 — N3 channel sideband gain-frequency characteristic.

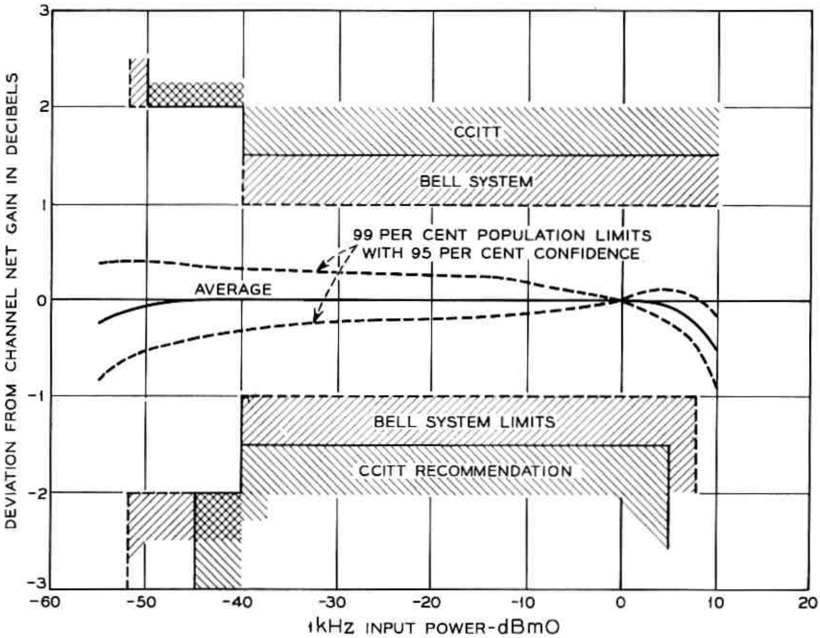

Fig. 7 — N3 channel compandor tracking characteristic.

The compandor tracking characteristics indicate the high degree of compatibility of the compressor and expandor circuits and the excellent reproducibility that is obtained with the manufactured product.

The N3 compandor circuit has a nominal 2:1 compression characteristic[3] and a 1:2 expansion characteristic. As a practical matter, the compressor and expandor circuits have been allowed to depart from the ideal 2:1 and 1:2 compression and expansion ratios. Fig. 8 presents the deviation from ideal 2:1 compression ratio as measured at the output of the compressor.

6.3 *Channel Distortion*

Table III summarizes the measured performance of the new N3 compandor units. The frequencies used in making these tests were 740 Hz and 1250 Hz. These frequencies were selected so that the important second- and third-order products would fall within the bandwidth of the voice-frequency channels, and none of the higher-order products would fall at any of the frequencies of the second- or third-order prod-

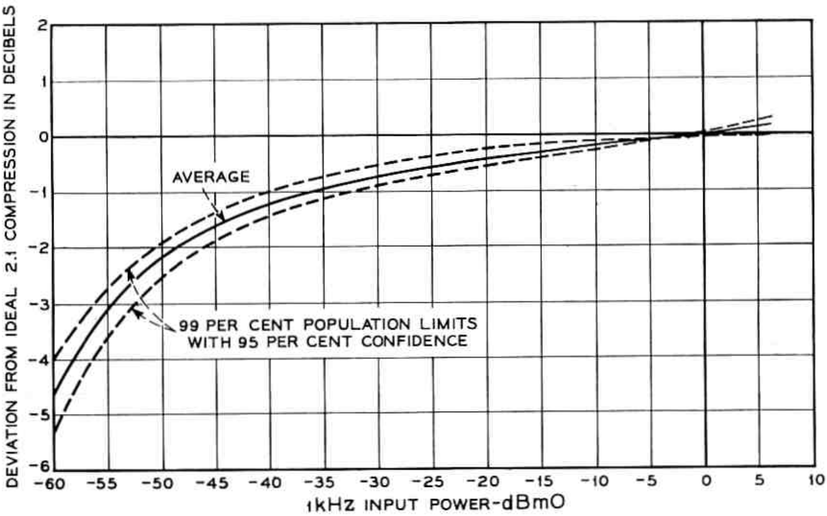

Fig. 8 — N3 channel compressor tracking characteristic.

ucts. The frequency differences between any two products, up to and including tenth order, or between fundamentals and modulation products is large enough that the desired modulation product can be measured with available test equipment.

Intramodulation products generated in the compressor and expandor circuits are about equal in magnitude. Their total magnitude will depend upon the phase relationships of the two products. If we assume that the value of the product is the same for both compressor and expandor, the maximum value will be for in phase products and have a magnitude 6 dB greater than the product from either circuit. If the separate products are 180 degrees out of phase, there will be complete cancellation and no product at the output of the compandor. The spread of maximum and minimum values of Table III reflect that distribution.

The above table gives the values of intramodulation distortion for all second- and third-order products. This has been done since there is considerable spread in the measured values for products of the same order but different frequencies. It is interesting to note that reduction of input signal level does not result in any substantial improvement in modulation distortion. The distortion certainly does not follow the power series law where a reduction of input power by 10 dB would result in a reduction of 10 dB in the relative level of second-order prod-

TABLE III — INTRACHANNEL DISTORTION FOR COMPANDORED N3 CHANNELS

Type	Frequency Hz	Value	Fundamental to Distortion Product Ratio—dB		
			Avg. Value of Fund. Power	Change in Ratio for Lower Fund. Power	
			0 dBmO	−10 dBmO	−20 dBmO
2a	1480	Avg. Max.* Min.*	47.7 40.5 56.1	2.3	0.9
2b	2500	Avg. Max. Min.	48.4 41.0 58.0	2.9	3.9
a + b	1990	Avg. Max. Min.	43.1 35.8 58.3	2.0	2.7
b − a	510	Avg. Max. Min.	42.7 35.4 57.0	−2.2	−5.0
3a	2220	Avg. Max. Min.	51.9 46.9 59.2	2.0	1.7
2a − b	230	Avg. Max. Min.	40.1 33.9 46.5	4.6	4.9
2b − a	1760	Avg. Max. Min.	37.1 34.8 39.5	2.4	2.9
2a + b	2730	Avg. Max. Min.	46.6 38.9 56.4	0.7	−0.6
2b + a	3240	Avg. Max. Min.	46.0 37.0 66.0	−0.5	−1.0

* The maximum and minimum values given in Table III are those obtained with the relatively small sample of 24 compandor units. The maximum values are indicative of the values expected. Tolerance limits are not given for this small sample since the summation of the distortion products developed in the over-all channel should not result in a normal distribution of values.

uct and 20 dB in the relative level of third-order product. The distortion products are generated in the variolossers where by design, the signal current-to-bias current ratio for the various input levels remains essentially constant; therefore, the level of the distortion products should remain essentially constant.

Table IV summarizes the performance of the compressor circuit alone as measured at the modulator input of an N3 channel.

6.4 *Channel Noise*

The back-to-back unloaded terminal noise performance of N3 channels is summarized in Table V. The performance of both compandored and non-compandored channels is well within the objectives.

TABLE IV — MODULATION DISTORTION OF THE N3 COMPRESSOR CIRCUIT

Type	Frequency Hz	Value	Fundamental to Distortion Product Ratio — dB		
			Avg. Value of Fund. Power	Change in Ratio for Lower Fund. Power	
			0 dBmO	−10 dBmO	−20 dBmO
2a	1480	Avg. Max. Min.	48.9 43.4 54.2	0.8	0.3
2b	2500	Avg. Max. Min.	50.2 43.2 56.8	1.6	0.9
a + b	1990	Avg. Max. Min.	46.2 40.2 54.7	0.0	−0.7
b − a	510	Avg. Max. Min.	43.5 36.9 47.8	−1.0	−3.5
3a	2220	Avg. Max. Min.	58.3 49.1 63.8	0.9	−0.3
2a − b	230	Avg. Max. Min.	41.9 36.8 47.5	0.4	−0.2
2b − a	1760	Avg. Max. Min.	44.9 40.6 49.1	−1.2	−1.5
2a + b	2730	Avg. Max. Min.	47.1 39.9 53.7	1.8	0.4
2b + a	3240	Avg. Max. Min.	44.8 38.8 50.0	2.6	1.6

TABLE V — UNLOADED TERMINAL CHANNEL NOISE

Type of Channel	Channel Noise at 0 TLP					
	dBrnC			dBrn 3 kHz Flat		
	Average	Max.	Min.	Average	Max.	Min.
Compandored	8.1	11.4	5.7	11.1	14.4	7.8
Non-Compandored	27.0	33.2	20.9	30.0	34.2	24.5

Measurements of the channel noise introduced by loading various combinations of channels with simulated speech indicated that certain types of system loading were detrimental to system performance. For example, several channels of one system may be loaded with identical signals by a few of the services offered by the operating companies. This has been termed "coherent" loading. Table VI gives the magnitude of noise generated by the loading of N3 channels with simulated speech, both coherent and non-coherent.

Loading the channels of an N3 carrier system with non-coherent noise does not impose a message noise problem as shown by the results when 22 channels are loaded. However, a noise advantage is obtained by loading certain channels when coherent loads are transmitted. The maximum reading for line (1) was reduced to 12.5 dB by also loading channels 2 and 6 of channel group 1 with the same signal for a total of ten loaded channels, emphasizing the benefit of the reversed phase of the modulating carriers of channels 2 and 6.

The noise performance of a carrier channel is judged on two bases: the C message weighted noise interference to a transmitted signal, and also the magnitude of impulse noise peaks. Large simultaneous amplitude peaks of speech in several channels of the N3 carrier system can

TABLE VI — CHANNEL NOISE DUE TO CHANNEL LOADING

No. of Chans Loaded	Type of Load	Channel Numbers	Grp No	Input Pwr at 0 TLP	Trsg Slope dB	Noise Value in dBrnCO	Message Rating
5	Coherent	Any combination	Any	0 Vu	0	<10	Satis
5	Non-coherent	Any combination	"	"	6	<10	"
(1) 8	Coherent	1, 3, 4, 5, 7, 8, 9, 11	1	"	0	21–26	Unsat
8	"	3, 4, 5, 7, 8, 9, 11, 12	2	"	0	26–33	"
8	"	3, 4, 5, 7, 8, 9, 11, 12	2	"	+6	35–40	"
8	"	1, 3, 4, 5, 7, 8, 9, 11	1	"	+6	≦10	Satis
10	Non-coherent	1 to 10 inclusive	1	"	0	≦10	"
22	"	All but chan 11, 12	1	−5 Vu	+6	≦10	"
22	"	All but chan 11, 12	1	−8 Vu	+6	≦10	"

generate impulse noise peaks in many channels by intermodulation processes within the group equipment. The input power of each of 22 channels loaded with simulated speech as disturbing signals had to be reduced from 0 Vu to −8 Vu at 0 TL to satisfy impulse noise objectives when the transmitting slope was +6 dB. This impulse noise performance is considered satisfactory.

The average measured noise advantage of the N3 expandor is 31.8 dB with a range between 30.8 and 32.9 dB. The subjective expandor advantage[8] is considered to average 5 dB less than the above figures when speech is being transmitted on the channel. Comparison of the unloaded noise performance of the two types of N3 channels shows an apparent expandor advantage of less than 20 dB. The unloaded noise of a compandored channel is usually controlled by the noise figure of the transistor at the input to the expandor amplifier (a low level point) which follows the expandor variolosser. Hence, the noise from this source is not reduced by the expandor loss.

6.5 *Intrasystem Crosstalk*

The far-end crosstalk performance of the N3 carrier terminals was one area where major improvements were incorporated, and the measured crosstalk values of the manufactured product reflect these improvements.

Near-end intrasystem crosstalk has never been a problem with the short-haul carrier systems since the carrier frequency allocations and operational methods were selected to eliminate the common sources of line near-end crosstalk at carrier frequencies. There were only a few measurable values of near-end intrasystem crosstalk observed with the N3 carrier systems, and these were well within the objective.

Only the minimum equal level coupling loss measured for each disturber is given in the results. Any other value would have little meaning since numerous combinations of disturbing and disturbed channels had no detectable crosstalk. In other combinations, crosstalk could be heard in the monitoring receiver but the noise of the disturbed channel masked the crosstalk contribution to the over-all meter reading. Crosstalk of this magnitude is not a problem. Far-end intrasystem modulation crosstalk generally had well-defined patterns for single-tone disturbers. However, it was never the controlling source, even for the lowest values of crosstalk interference measured.

Table VII gives the minimum equal level crosstalk coupling losses measured at the expandor output.

The out-of-band suppression of the channel filters can best be de-

TABLE VII — MINIMUM FAR-END EQUAL LEVEL CROSSTALK COUPLING LOSS AT EXPANDOR OUTPUT

Disturbing Frequency (0 dBmO)	Loss — dB	
	"C Mess Wtg"	"3 kHz Flat Wtg"
200 Hz	70.1	66.5
1000 Hz	77.4	77.7
3450 Hz	83.1	76.2
5000 Hz	86.4	83.6
Simulated Speech	80.9	75.7

termined by measuring the higher level crosstalk at the output of the demodulator. The crosstalk-to-noise ratio at this point for a 0 dBmO disturber is sufficiently large to assure accurate crosstalk measurement. Table VIII gives the minimum equal level coupling losses measured at the demodulator output.

6.6 *Net Loss Stability*

Stability of transmission has generally been referred to in terms of net loss of the derived trunk. The term net loss has been used in this section for clarity of association. The net loss stability of the N3 channels has been determined by measuring the gain at 1000 Hz at frequent intervals over a period of several months. Net loss stability tests are in progress on several operating N3 carrier systems of about average length. Table IX summarizes the measurements that have been made to date.

No adjustments have been made on any of these channels during the test period. The results are well within design objectives.

TABLE VIII — MINIMUM FAR-END EQUAL LEVEL CROSSTALK COUPLING LOSS AT DEMODULATOR OUTPUT

Disturbing Frequency (0 dBmO)	Loss — dB	
	"C Mess Wtg"	"3 kHz Flat Wtg"
200 Hz	41.7	38.4*
1000 Hz	43.9	46.9
3450 Hz	49.3	45.7
5000 Hz	55.9	55.3
Simulated speech	48.0	48.0

* All of these equal level coupling losses satisfy the objectives with the exception of the minimum value for 3 kHz flat weighting at 200 Hz. Only this single measurement fell short of the 40 dB objective. The average value for a sample of 48 measurements was 45.5 dB.

TABLE IX — N3 CHANNEL NET LOSS STABILITY

System	No. of Channels	Net Loss Stability			
		3 Mos.		6 Mos.	
		Bias (Avg)	Dist gr (Std Dev)	Bias (Avg)	Dist gr (Std Dev)
Hartford-Waterbury	96	−0.04	0.16	+0.20*	0.27*
Dayton-Springfield	78	−0.03	0.13	−0.094	0.20

* The elapsed time between initial and these measurements was seven months.

6.7 *Envelope Delay*

Fig. 9 shows the envelope delay distortion of N3 compandored channels as measured with terminals connected back-to-back. This distortion is due primarily to the channel filters. The type of voice-frequency unit (compandor or VF amplifier) has only a small effect on the magnitude of the distortion. The curve of Fig. 9 is applicable for both types of channels.

The absolute delay of a carrier channel is also important for some installations. Table X gives the absolute terminal delay in microseconds applicable to channels equipped with either the compandor or VF amplifier units. The absolute delay of a typical N-repeatered line is about 10 microseconds per mile.

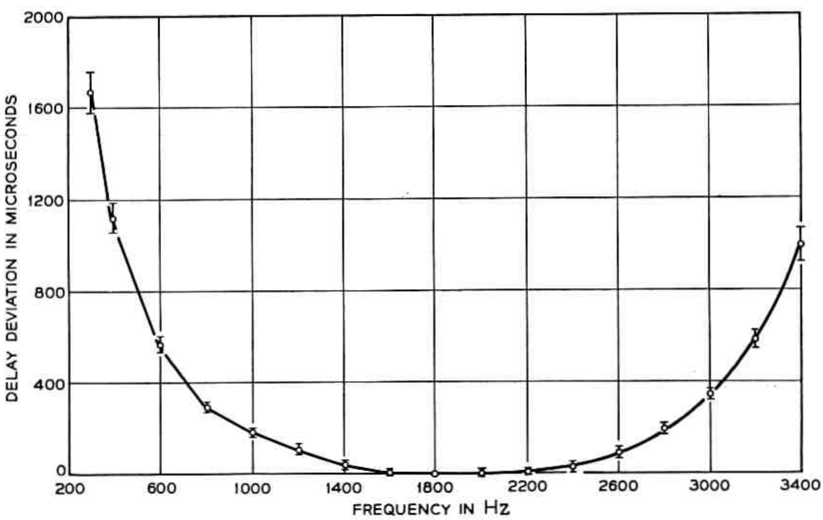

Fig. 9 — N3 channel delay distortion characteristic.

TABLE X — N3 CHANNEL DELAY AT 1800 Hz

Input Power dBmO	Absolute Terminal Delay in Microseconds	
	Avg. Delay	Limits for 99% of Chans with 95% Confidence
+1.0	983	±78.0
−9.0	1001	±76.0
−14.0	1006	±74.5

6.8 *Alarm Operation*

Tests of the carrier failure alarms, automatic trunk conditioning and automatic service restoral features of the N3 carrier system have indicated satisfactory performance within design objectives. Measured alarm thresholds (loss in carrier power sufficient for alarm registration) ranged from 17 to 20 dB below nominal received carrier power. Such threshold values had been established to assure satisfactory operation during the restoral gain transient of an N-repeatered line.

Idle channel noise at the instant of service restoral ranged from 33.5 to 37.0 dBrnCO. While this noise value is above the normal system objective, experience has shown that any restored N3 carrier system whose channel noise measurements are in the above range can meet the severe system objectives within a short time. This prevents restoral of systems with degraded transmission performance.

VII. ACKNOWLEDGMENT

The development of a carrier telephone system necessarily involves the contributions of many people. While it is not possible to mention all, the authors wish to acknowledge, with thanks, the contributions of F. H. Blecher, R. C. Boyd, F. J. Herr, Mrs. R. A. King, W. R. Lundry, D. D. Sagaser, T. W. Thatcher, and L. F. Willey.

REFERENCES

1. Caruthers, R. S., The Type N1 Carrier Telephone System: Objectives and Transmission Features, B.S.T.J., *30,* Jan., 1951, pp. 1–32.
2. Fracassi, R. D. and Kahl, H., Type ON Carrier Telephone, Trans. AIEE, *72,* Jan., 1954, pp. 713–21.
3. Boyd, R. C. and Herr, F. J., The N2 Carrier Terminal — Objectives and Analysis, B.S.T.J., *44,* May–June, 1965, pp. 731–59.
4. Haner, R. L. and Wood, I. E., N3 Carrier System — Circuit Design, B.S.T.J., This Issue, pp. 801–844.

5. Bowie, L. A., Giguere, W. J., and Watters, J. A., Equipment Design of N3 Carrier Terminals and Carrier Supply, B.S.T.J., This Issue, pp. 845–894.
6. Weaver, A. and Newell, N. A., In-Band Single Frequency Signaling, B.S.T.J., *33*, November, 1954, p. 1381.
7. Lundry, W. R. and Willey, L. F., The N2 Carrier Terminal — Circuit Design, B.S.T.J., *44*, May–June 1965, pp. 761–87.
8. Carter, C. W., Jr., Dickieson, A. C., and Mitchell, D., Application of Compandors to Telephone Circuits, AIEE, *65*, 1946, pp. 1079–1086.

Circuit Design of the N3 Carrier Terminal

By R. L. HANER and I. E. WOOD

(Manuscript received March 17, 1966)

In the circuit design of the N3 carrier terminal full use has been made of the advances in solid state art, crystal filter design techniques, and miniature ferrite transformer performance. Economies have been provided in power consumption, space requirements, and maintenance effort. Performance has been significantly improved over its predecessor the ON2 terminal in the areas of channel response, net loss stability, compandor tracking, and modulation distortion. Two concepts, new to the short-haul carrier field, have been employed in the N3 design. Frequency correction circuits are used to precisely correct for frequency shifts accumulated as the signal is transmitted over the N-repeatered line. A common carrier supply is utilized to provide extremely precise carriers economically for terminal use.

I. INTRODUCTION

During the late 1950's short-haul carrier systems began to undergo a comprehensive modernization program. The primary objective of this program was to achieve significant improvement in reliability and transmission performance of the short-haul family in order to meet higher standards for transmission of message, program, and data or other special services over intertoll trunks. Equally desirable improvements were sought in the areas of miniaturization, reduced maintenance, economical installation, lower power requirements, and flexibility. The first two phases of this program have been completed with the introduction of the N1A repeater and the N2 terminal — a twelve-channel double-sideband system. The third phase, the N3 carrier terminal, is described from a circuit standpoint in this paper.

The N3 carrier terminal, a new 24-channel single-sideband system, was designed to provide a considerable transmission improvement over its predecessor the ON2 system. It has been engineered to provide a 3-dB channel bandwidth between 200 and 3450 Hz and a net loss stability of ±0.5 dB over extreme operating conditions. A large part of the performance improvements[1] have been made possible by the rapid growth of the solid-state device art. With the N3 terminal, several new system concepts have appeared for the first time in the short-haul family. A common carrier supply for many systems as opposed to individual per channel oscillators is utilized. This provides for the economical generation of many extremely precise carriers for terminal use. A second new concept is the frequency correction circuit which removes frequency shifts accumulated along the repeatered line. Improvements in both crystal filter design and ferrite core transformers have also played an important role in obtaining the superior performance in the N3 terminal.

II. CARRIER TERMINAL

2.1 *Transmitting Terminal*

A block diagram indicating frequencies within a transmitting terminal is shown in Fig. 1. The input voice-frequency signal is first compressed and then applied to a channel modulator circuit. The channel modulator translates the signal, selecting the upper sideband, into a 4-kHz slot in the 148 to 196-kHz range. The twelve resulting outputs when combined in the resistive combining multiple form a solid lay-up of signals in this spectrum. In addition, six transmitted carriers (152, 160, 168, 176, 184, and 192 kHz) are combined at a precise level with the signal in the multiple. This composite signal is translated a second time to the 36 to 84-kHz band by the channel group modulator. A second composite signal is similarly translated to the 84 to 132-kHz band by a second channel group modulator. The outputs of these two circuits are combined in a hybrid coil network to form a low group signal consisting of 24 4-kHz channels and 12 carriers. Depending on the type of group units used, this signal can be transmitted in either the 36 to 132-kHz or the 172 to 268-kHz band. The line terminating circuit is used to interconnect the group circuits and the carrier line. The circuit provides plug-in span pads to build out the loss of short cable sections. Secondary lightning protection is incorporated to limit surges to approximately 30 volts. Finally, the line terminating circuit provides a flexible means of simplexing power over the line to energize up to three remote transistorized N1A repeaters or one electron tube N1 repeater.

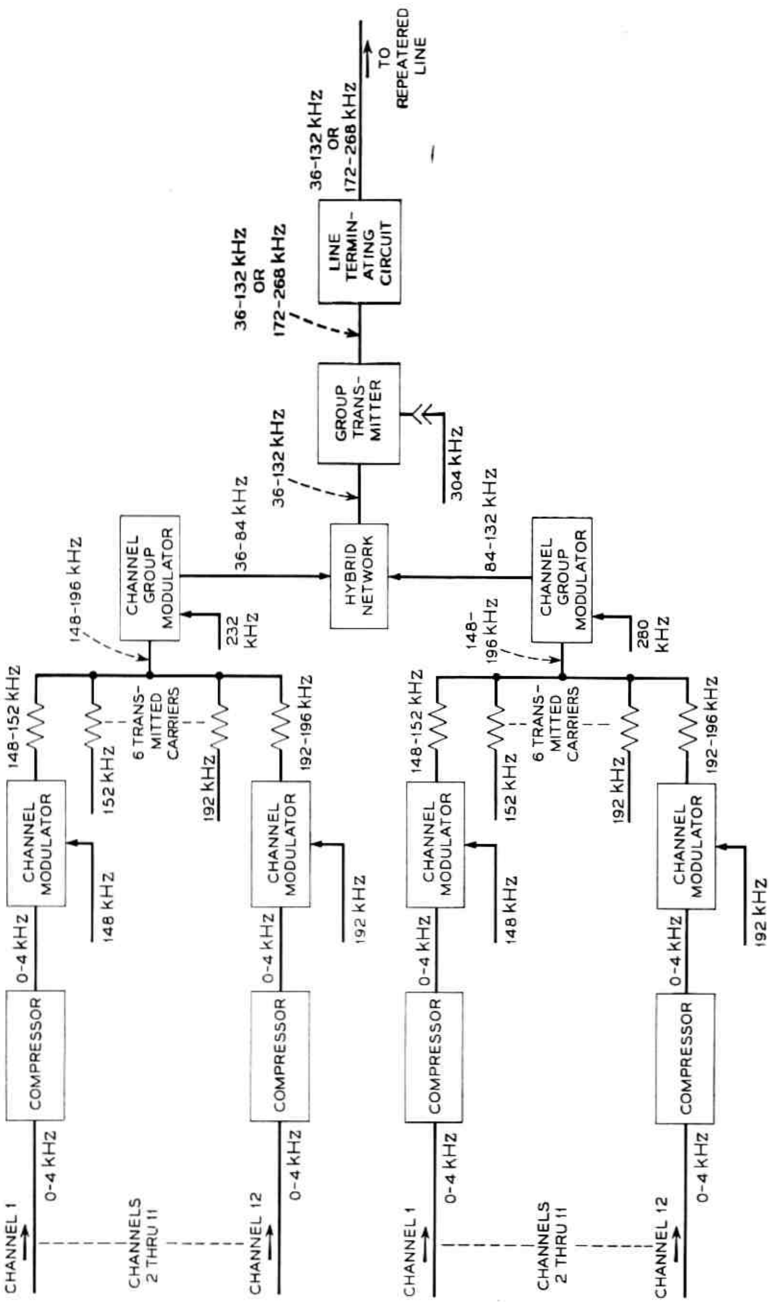

Fig. 1 — N3 transmitting terminal.

2.2 *Receiving Terminal*

A block diagram of the receiving terminal is shown in Fig. 2. The received signal is applied to either a low group or high group receiver circuit. The output is transmitted to the two channel group demodulators via a resistive splitting network. Each of the channel group demodulators selects and translates half of the low group band to the 148 to 196-kHz range. The appropriate carrier for demodulation is derived in the frequency correction circuit in such a fashion as to correct for any frequency shift in the incoming signal. The output from the channel group demodulator is applied to six double-channel regulator circuits. Each of the regulators dynamically compensates for over-all system gain variations on a pair of channels by regulating the carrier pilot between them to a precise level. Each of the dual outputs of the regulator circuit is applied to a channel demodulator circuit which contains a crystal channel band filter to separate a particular 4-kHz signal from the 148 to 196-kHz band. After demodulation to voice frequency the signal is applied to the expandor circuit which restores the original volume range.

A 48 to 21-volt power converter is used for powering each terminal. Since present types of transistors cannot efficiently operate with supply voltages much greater than 20 volts, approximately 50 per cent reduction in power dissipation is realized by using the converter. In addition, the converter contains a regulator circuit which provides an extremely well regulated voltage that is practically independent of wide fluctuations in the −48 volt central office battery. This feature essentially eliminates dependence of the net loss stability performance of the system on the office battery voltage variations. An alarm feature is also included to provide an indication when the converter output is out of limits.

III. COMPANDOR CIRCUIT

The companding technique is utilized to gain a noise and crosstalk advantage to achieve satisfactory transmission performance using the economical N-type repeatered lines. The compandor consists of two complementary circuits: a compressor in the transmitting terminal and an expandor in the receiving terminal.

The compressor reduces the range of signal volumes at its input by a factor of two. This is accomplished by a diode shunt variolosser followed by a constant gain three-stage feedback amplifier. A portion of the output of the amplifier is rectified and used as the control

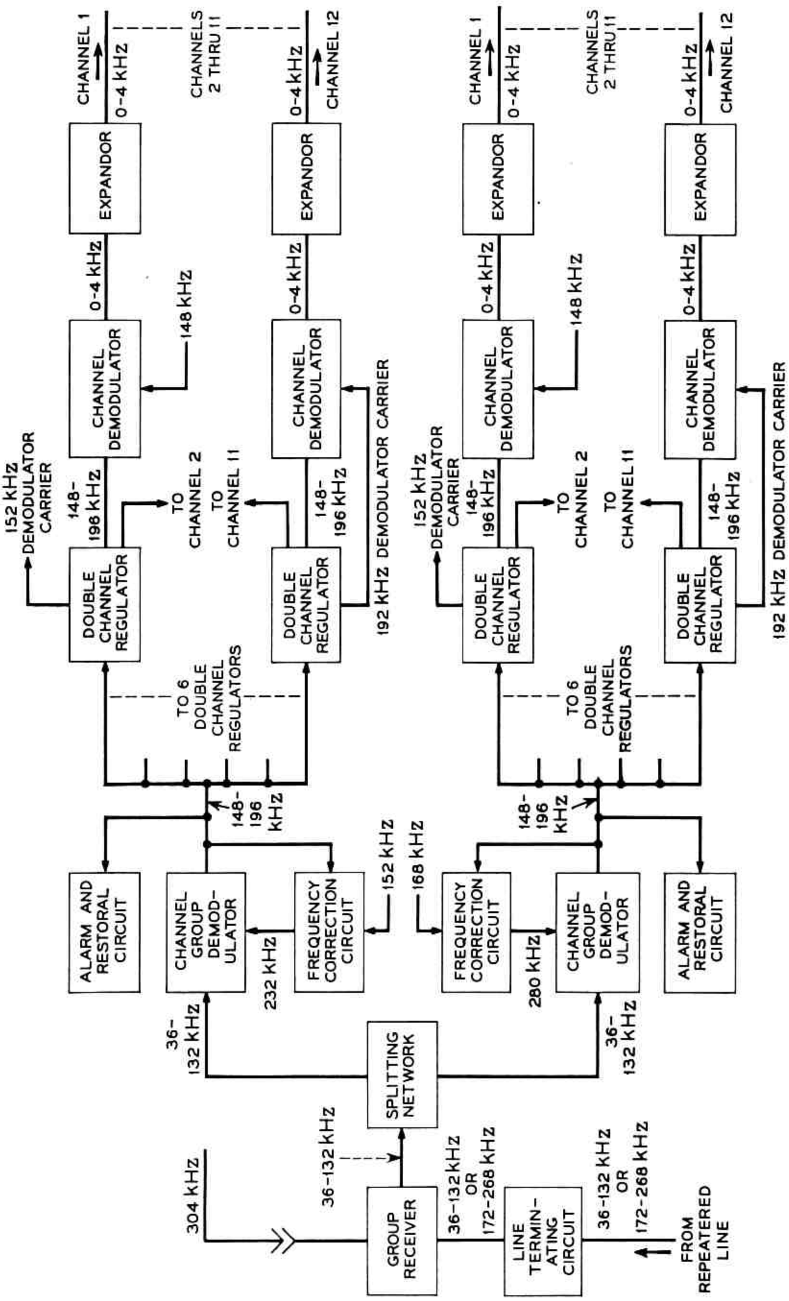

Fig. 2 — N3 receiving terminal.

signal to the variolosser. The action is such that the stronger signals are relatively unchanged while the weaker tones are raised in level. Thus, for a 60-dB input range, the output varies by only 30 dB.

The expandor increases the range of signal volumes that it receives at its input by a factor of two. This is accomplished by a series diode variolosser followed by a three-stage feedback amplifier. A portion of the input signal is amplified in a control amplifier, rectified, and used as the control signal to the variolosser. The action is such that the stronger signals are relatively unchanged while the weaker tones are lowered in level. Thus, for a 30-dB input range the output varies by 60 dB.

When a compressor in the transmitting terminal is followed by an expandor in the receiving terminal, the original volume range is restored with a tracking error which is typically less than 0.2 dB. The compandor used in the N2 carrier system, which from a circuit standpoint is identical to the N3 compandor, is described in detail in another article.[2] It is noteworthy that the expandor contains the only channel gain adjustment used for normal lineup and maintenance.

IV. CHANNEL MODEM CIRCUIT

4.1 *General*

The channel modem circuit consists of both the transmitting channel modulator circuit and the receiving channel demodulator circuit, both of which are packaged in a single plug-in unit. The channel modulator circuit provides the first step of modulation from voice to a 4-kHz portion of the 148 to 196-kHz band. The channel demodulator circuit provides the last step of demodulation from a 4-kHz portion of the 148 to 196-kHz band to voice frequency. Each of the twenty-four channel modem circuits in a terminal are identical. The various channel frequencies are selected by inserting the appropriate crystal channel bandpass filters and making the proper adjustment on the variable equalizer.

4.2 *Channel Modulator Circuit*

The channel modulator circuit shown in Fig. 3 consists of a temperature compensating input pad, a transistor switch-type modulator, a 6-dB isolating pad, and a crystal channel bandpass filter. The carrier drive for the modulator is supplied externally from the common carrier supply. The input pad contains a thermistor as one of the ele-

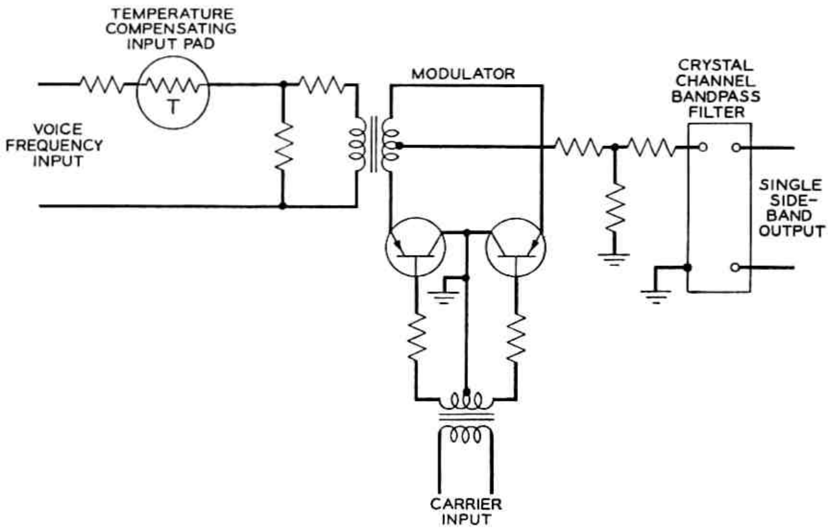

Fig. 3 — Channel modulator circuit.

ments such that the pad loss varies inversely with temperature and compensates for the entire modulator circuit. The transistor used in the modulator represents the only new semiconductor device developed especially for N3. This device consists of a matched pair of germanium alloy transistors with an inverse beta of at least 15 (inverse beta is the current gain when the collector is used as the emitter). The high inverse beta is required to minimize the carrier drive necessary for satisfactory operation of the modulator. The transistors are matched to provide sufficient carrier balance at the output of the modulator.

The modulator operation is most easily visualized by noting that the carrier controls the transistor switching; one transistor is "on" and the other is "off" for a given polarity of the carrier. The transistor operation is reversed for the opposite polarity of the carrier. With the transistors switching "on" and "off" in this way, the output termination is switched from one half of the secondary winding of the input transformer to the other at carrier frequency. Therefore, at any instant of time, one half of the secondary is properly terminated and the other is open-circuited. This operation reverses the polarity of the incoming signal at a carrier frequency rate and thus produces a double-sideband suppressed-carrier signal at the output of the modulator.

The resistors in the base leads provide a constant current source and a good carrier drive input impedance. For an input signal level of -15 dBm and a carrier drive of -6 dBm, the carrier leak at the output is at least 32 dB down on each of the sidebands. The $2A - B$ distortion product is approximately 48 dB down on either the A or B signal (each at -15 dBm). The modulator remains "linear" for input signals up to -11 dBm which represent $+8$ dBm at OTL. Following this modulator is a 6-dB pad to provide isolation between the modulator and crystal channel band filter. The characteristics of the channel filter which employs two self-equalized quartz crystal filter sections are shown in Fig. 4. Each of the twelve different filters was designed to pass the upper sideband (200 to 3450 Hz above the carrier frequency) in the 148 to 196 kHz range. The discrimination against the lower sideband is at least 55 dB while 30 dB of discrimination is provided 4 Hz above the carrier frequency. The in-band ripple is held to less than ± 0.15 dB at 80°F and less than ± 0.25 dB at 120°F.

4.3 *Channel Demodulator Circuit*

The channel demodulator circuit shown in Fig. 5 consists of a crystal channel bandpass filter, a 6-dB isolating pad on each side of the demodulator, a transistor switch type demodulator, a low-pass filter,

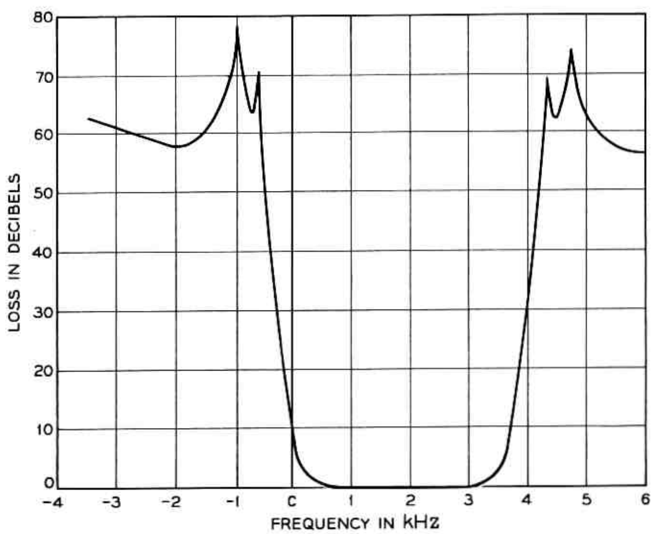

Fig. 4 — Measured insertion loss N3 channel band filter.

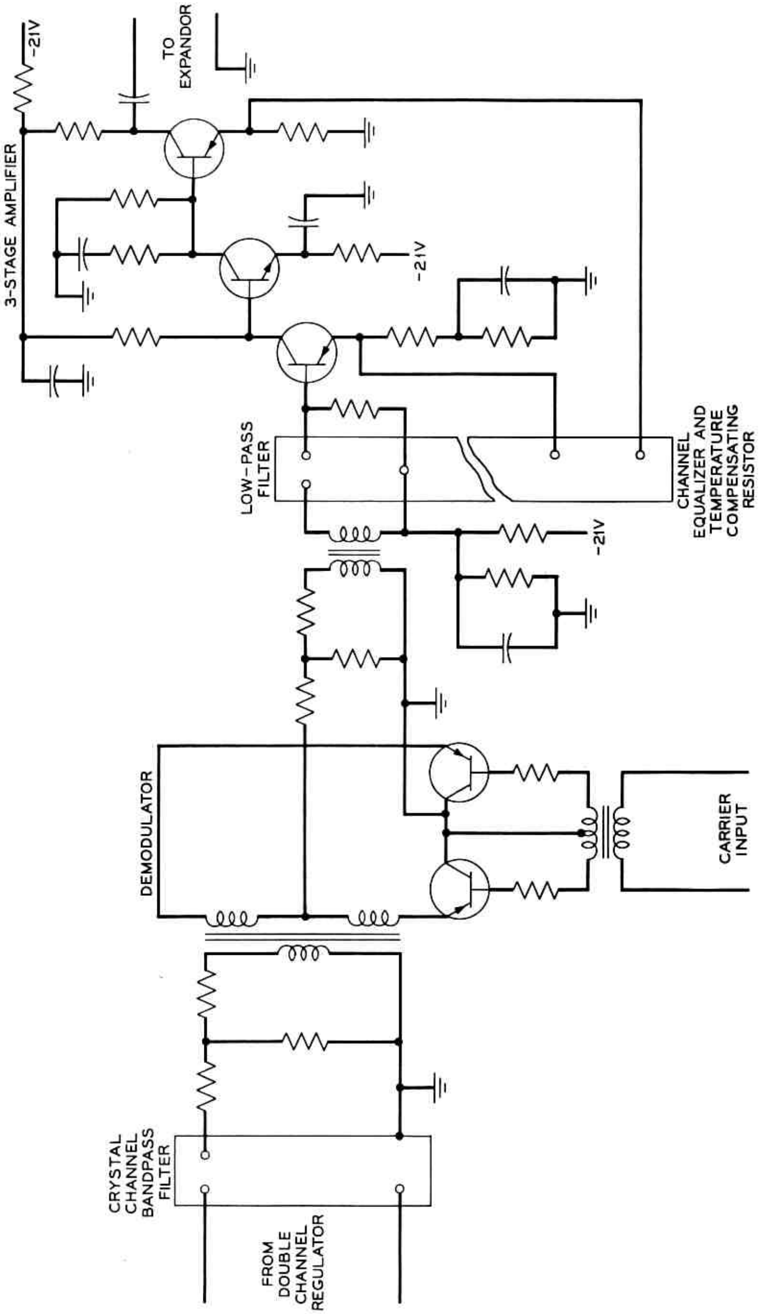

Fig. 5 — Channel demodulator circuit.

and a feedback amplifier. The carrier drive for the demodulator is supplied externally from either the common carrier supply (odd numbered channels) or the double channel regulator circuit (even numbered channels). The crystal channel bandpass filter, which is identical to that used in the channel modulator circuit, selects the proper 4-kHz sideband from the group of twelve present at the output of the double channel regulator. Following the filter is a 6-dB pad to provide isolation between the channel filter and the demodulator. The transistor demodulator translates the 4-kHz sideband down to the audio range. The low-pass filter, separated from the demodulator by a 6-dB pad, passes the audio band and rejects the higher frequency energy generated in the demodulator. A loss peak as shown in Fig. 6 is placed at 4 kHz to suppress any adjacent carrier energy present. The three-stage feedback amplifier following the low-pass filter includes as part of the feedback circuit a temperature compensating resistor and a channel equalizer. The positive coefficient resistor compensates for the entire demodulator circuit. The equalizer provides loss peaks which are transformed by the amplifier to gain peaks at approximately 120 Hz and 3650 Hz to equalize for the roll-off of both the transmitting and receiving channel band filters. The characteristics of the equalizer when operating in the feedback circuit of the amplifier are shown in Fig. 7. The position and shape of these peaks may be adjusted slightly by means of screw-down options to accommodate small differences in the various band filters for different channels. The open loop gain of the amplifier at midband (1000 Hz) is approximately 24 dB. At the band edges (200 and 3450 Hz) the feedback is reduced to approximately 22 dB due to the loss peaks in the equalizer. The open loop gain cut-off is about 1 MHz with more than 55 degrees phase margin to insure stable operation. The low-pass filter, equalizer, and temperature compensating resistor R_T are all pack-

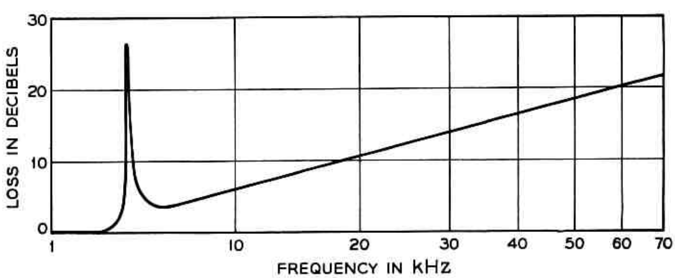

Fig. 6 — Low-pass filter.

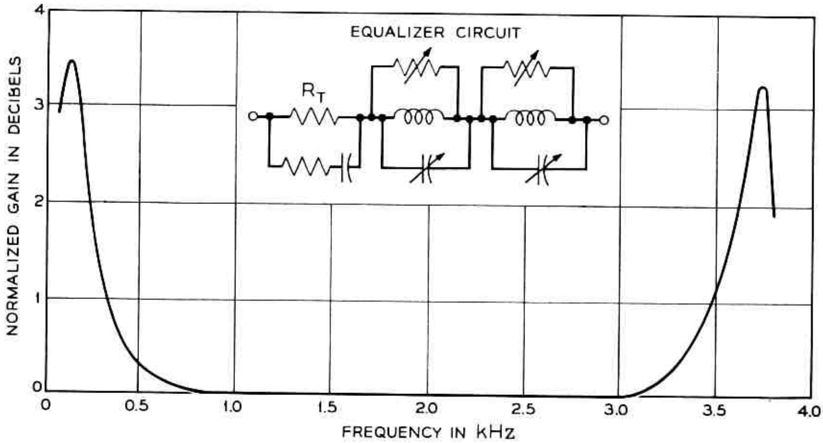

Fig. 7 — Frequency characteristic of demodulator amplifier including channel equalizer.

aged as a single piece of apparatus. For an output signal level of −5 dBm, which represents 0 dBm at OTL, the 2A − B distortion product due to the entire demodulator circuit is approximately 48 dB down on either the A or B signal. The circuit remains linear for output signals up to −1 dBm.

V. DOUBLE-CHANNEL REGULATOR CIRCUIT

Since the carrier to reference level sideband ratio is precisely established and maintained in the transmitting terminal, the carrier tone may be used in the receiving terminal as a pilot to provide regulation. The function of the double-channel regulator circuits (AGC circuit) is to automatically regulate each pair of channels by using the received carrier between them as a pilot tone. Basically the regulator is an amplifier, the gain of which is inversely proportional to the input level of a pilot carrier. Each of the twelve double-channel regulator circuits in a terminal is identical. The specific pair of channels to be regulated is selected by inserting a particular crystal carrier pick-off filter in the regulator unit. The input signal to the double-channel regulator circuit consists of twelve single-sideband 4-Hz signals and six carrier tones. Although only two of the signals and the carrier between them are of interest insofar as regulation is concerned, the additional carriers and signals which are present impose a severe intermodulation distortion requirement on the circuit. In addition to providing the regulating function, the regulator supplies one of the adjacent channels (even numbered channel) with a channel demodulating carrier. Since this carrier is derived from the carrier supply at the transmitting terminal, the process provides for error-free voice frequency recovery in all the even numbered channels.

The transmission path for the double-channel regulator circuit shown in Fig. 8 consists of a shunt variolosser and a 4-stage forward transmission amplifier. Negative feedback around the last three

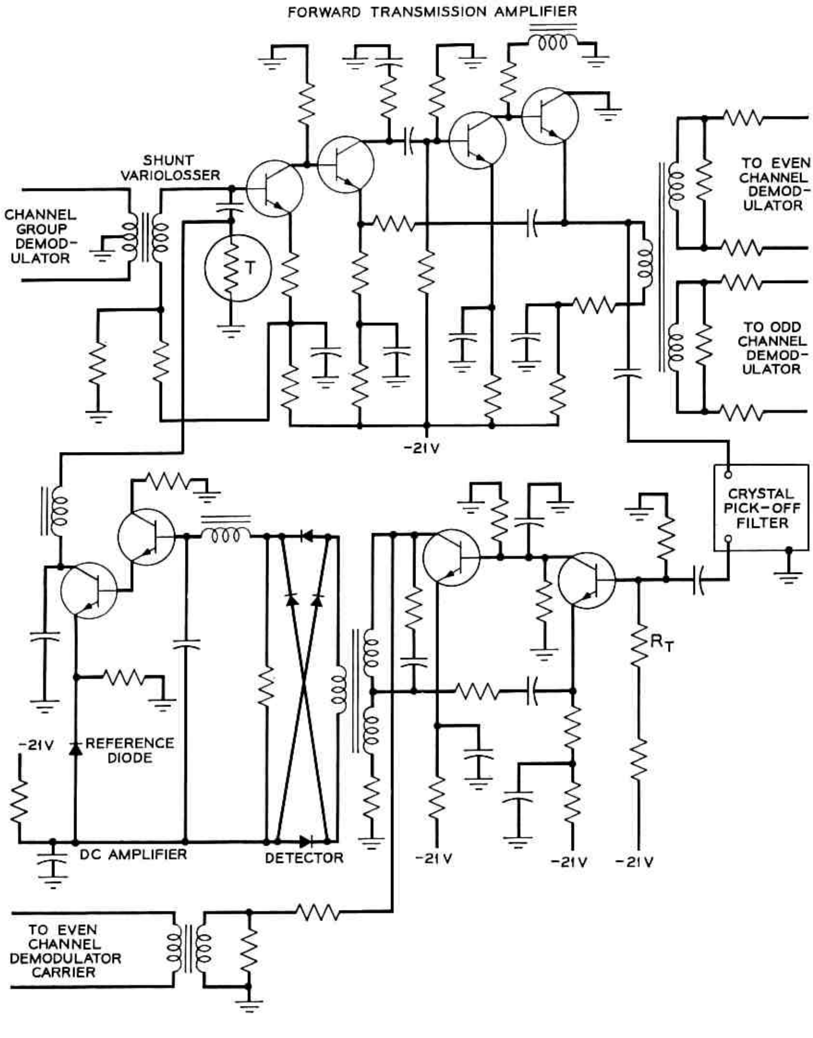

Fig. 8 — Double-channel regulator.

stages, in addition to local feedback on the output stage is employed to provide the necessary intermodulation distortion performance. At the normal operating level for a 0 dBm at OTL tone the 2A − B distortion product is approximately 112 dB down on either the A or B signal. A plot of the feedback ($\mu\beta$) phase and gain characteristics for the forward transmission amplifier is shown in Fig. 9. The high end gain cut-off occurs at approximately 10 MHz with a phase margin of 50 degrees. The low end gain cut-off occurs at 860 Hz. By using the following relationship, the phase margin at the low end was calculated to be approximately 60 degrees:

$$\left|\frac{\mu}{1-\mu\beta}\right| \Big/ \left|\mu\right| = \left|\frac{1}{1-\mu\beta}\right| = \frac{1}{2\sin\frac{\theta}{2}}$$

where

θ = phase margin

$|\mu|$ = magnitude of μ gain at crossover frequency

$\left|\frac{\mu}{1-\mu\beta}\right|$ = magnitude of closed loop gain at crossover frequency.

The transformer output with a split secondary is used to provide

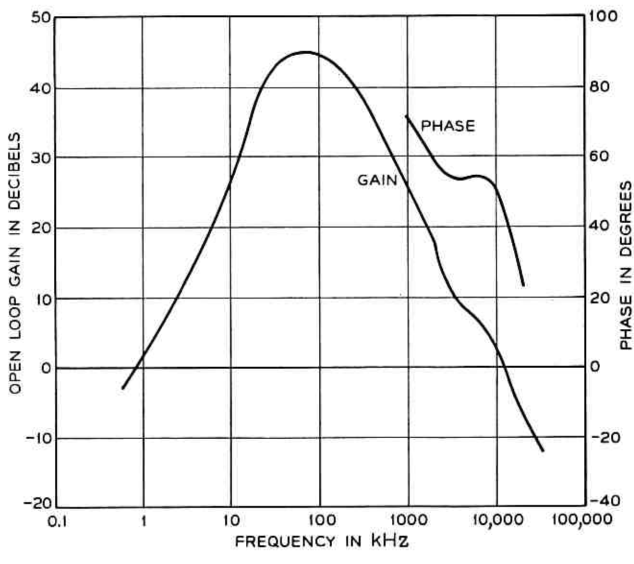

Fig. 9 — Open loop phase and gain of forward transmission amplifier.

separate outputs to the odd and even channel demodulators. By using an emitter-follower output stage and shunt feedback, a high degree of isolation is achieved between the two outputs. The small resistor in series with the primary of the output transformer reduces the isolation somewhat but is advantageous in improving the intermodulation distortion performance of the transformer. If one demodulator is removed, the level in the other will change by less than 0.1 dB. The low output impedance of the amplifier also permits a lower bridging loss of the crystal pick-off filter consistent with the ripple objective of 0.05 dB maximum in adjacent channels. Because of the frequency correction technique employed in the receiving terminal, the carrier pick-off filter can be made extremely narrow and thus achieve a high discrimination to all other frequencies. The carrier level at the output of the pick-off filter is amplified, rectified, filtered, and compared to the reference diode voltage. The difference or error voltage is amplified and used as the control current in the thermistor variolosser. Thus, if the regulator output level increases, the error voltage increases, the thermistor impedance decreases, and therefore the loss in the variolosser increases tending to reduce the output level. The regulation characteristic shown in Fig. 10 indicates approximately a 0.3-dB output change for a 30-dB input change. The over-all or envelope open loop gain characteristic is shown in Fig. 11. This characteristic was obtained by breaking the loop at the input to the dc amplifier

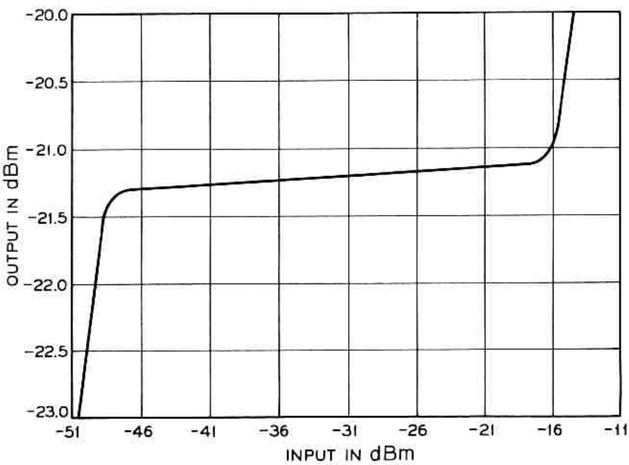

Fig. 10 — Regulation characteristic.

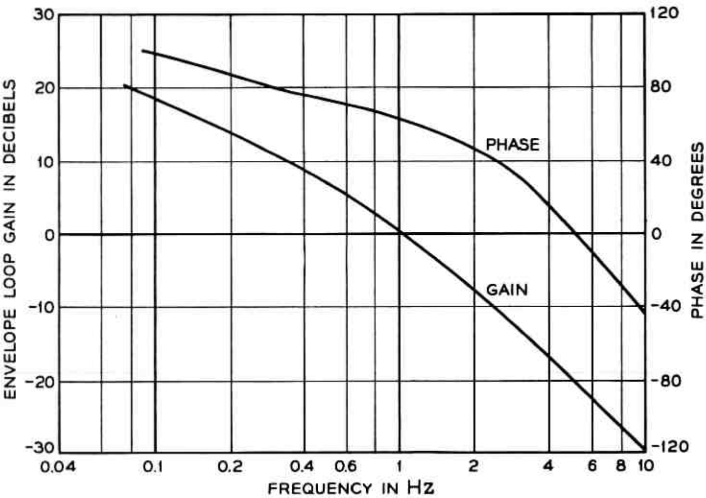

Fig. 11 — Envelope open loop phase and gain characteristic of double channel regulator.

and measuring the low frequency gain and phase. The high end cut-off occurs at approximately 1 Hz with a phase margin of 63 degrees. This extremely low cut-off frequency is very desirable so that the regulator will tend to discriminate against and not regulate or follow, carrier beats.

Temperature compensation of the entire regulator circuit is accomplished by means of a positive temperature coefficient shunt resistor R_T following the pick-off filter. A tap at the output of the two stage amplifier is used to provide the even channel demodulating carrier.

VI. CHANNEL GROUP MODEM CIRCUIT

6.1 *Channel Group Modulator*

In the transmitting terminal, the input signal to the channel group modem consists of 12 single-sideband signals plus 6 transmitted carriers in the 148 to 196-kHz range. The function of the channel group modulator is to translate this signal to either the 36 to 84 or 84 to 132-kHz ranges using a frequency of 232 or 280 kHz supplied from the common carrier supply. The channel group modulator circuit consists of a diode ring modulator and a single stage carrier drive amplifier.

The amplifier raises the incoming carrier by 7 dB and provides optimum impedances at the input and output for the carrier supply and modulator. The outputs of two channel group modulators are subsequently combined to give a composite signal in the 36 to 132-kHz band containing 24 sidebands and 12 carriers before being applied to the group transmitter.

6.2 *Channel Group Demodulator*

In the receiving terminal, the output of the group receiver is applied to two channel group demodulator circuits via a resistive splitting pad. Each half (36 to 84 or 84 to 132 kHz) of the low group signal is selected and demodulated into the 148 to 196-kHz band. This signal is then amplified for transmission to the double-channel regulator circuits. As shown in Fig. 12, the input signal to the channel group demodulator circuit is filtered to select the proper half of the low-group signal before being demodulated. The demodulator consists of a double-balanced diode ring-type demodulator. The single stage carrier driver amplifier provides gain, isolation, and optimum input and output impedances between the carrier source and the demodulator. The low-pass filter following the demodulator passes the 148 to 196-kHz signal and rejects the higher frequencies generated by the demodulator. Approximately 25 dB of negative feedback is incorporated in the 2-stage hybrid-feedback amplifier to stabilize the gain and provide an adequate intermodulation performance. When the modulator and demodulator circuits are connected in tandem the frequency characteristic obtained is shown in Fig. 13.

VII. FREQUENCY CORRECTION CIRCUIT

The signal entering an N3 receiving terminal, after being transmitted over an N-repeatered line, will have been shifted in frequency: a large shift being in the order of 100 Hz. This shift is due to an accumulation of small errors present in each of the 304-kHz oscillators in the frogging repeaters along the line. A frequency shift of this magnitude, if not corrected, would greatly deteriorate the over-all N3 performance. Since voice-frequency equalizers are used to equalize the extremely sharp carrier frequency crystal band filters, a frequency shift would seriously affect the channel frequency response. In addition, the frequency error that would remain in those channels using local carriers for demodulation would be intolerable.

The function of the frequency correction circuit is to introduce a

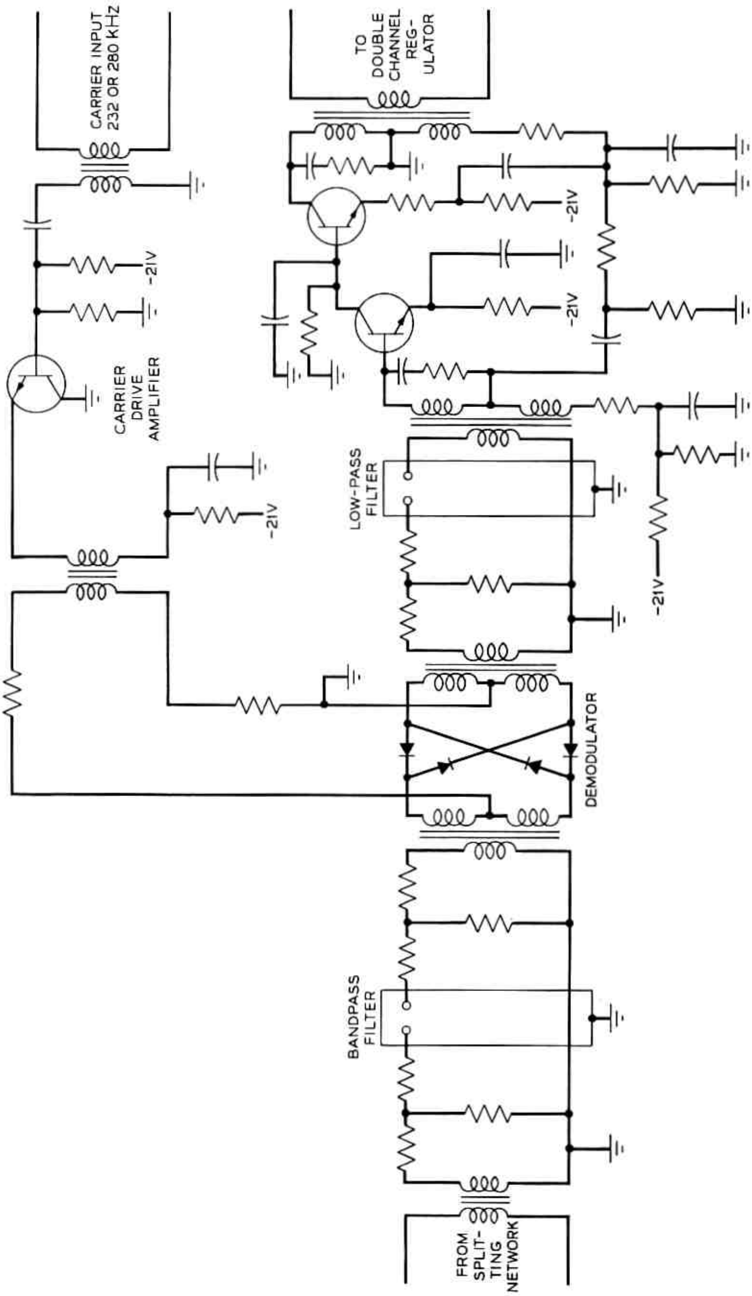

Fig. 12 — Channel group demodulator.

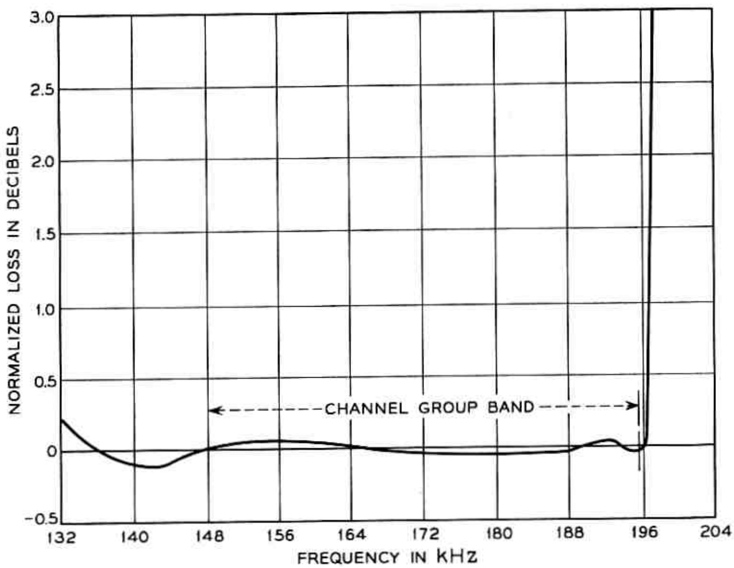

Fig. 13 — Frequency characteristic of channel group modulator and demodulator in tandem.

compensating frequency shift into the received signal so as to eliminate any frequency error accumulated along the line. In addition, this circuit substantially reduces the effects of frequency differences between the transmitting and receiving common carrier supplies. Two tones or carriers are required for this scheme: one, considered to be the reference, is derived from the receiving carrier supply and the other from the incoming signal.

The frequency correction circuit basically works on the phase-locked-loop principle.* The customary phase-locked loop has a severe disadvantage for applications of this type. It requires clearing the frequency spectrum around the carrier tone with which the loop works. A fundamental incompatibility exists between a wide capture range on one hand and high discrimination against nearby signals on the other. The frequency correction circuit overcomes this difficulty by incorporating two conventional loops and a switching control circuit to select the appropriate loop. One of the loops has the characteristic of capturing over a wide range of frequencies; the other is an extremely narrow

* Early N3 terminals used a forward acting frequency correction circuit instead of the technique described in this section.

band circuit which offers high discrimination to tones located close to the wanted carrier in the frequency spectrum. The function of the switching circuit is to switch in the wideband loop when the wanted carrier is sufficiently off-frequency, and then the narrow-band loop after a frequency lock has been achieved.

The block diagram shown in Fig. 14 will help to illustrate the operation of the frequency correction circuit. An incoming signal to the channel group demodulator consists of twelve single-sideband and six transmitted carriers in the 84 to 132-kHz portion of the spectrum. As indicated on the diagram this received signal has been shifted by some amount Δ during the course of its transmission over the N-repeatered line. The function of the channel group demodulator is to transform this signal into the 148–196 kHz range. This is accomplished by supplying a nominal 280-kHz carrier and selecting the lower sideband output. Initially, this output band may be shifted in frequency. During this initial period switches S1 and S2 are in the (a) position. A bridging amplifier connected to the output of the channel group demodulator amplifies this shifted signal which is applied to the pick-off filter and bypass pad. The amplifier contains an LC resonant

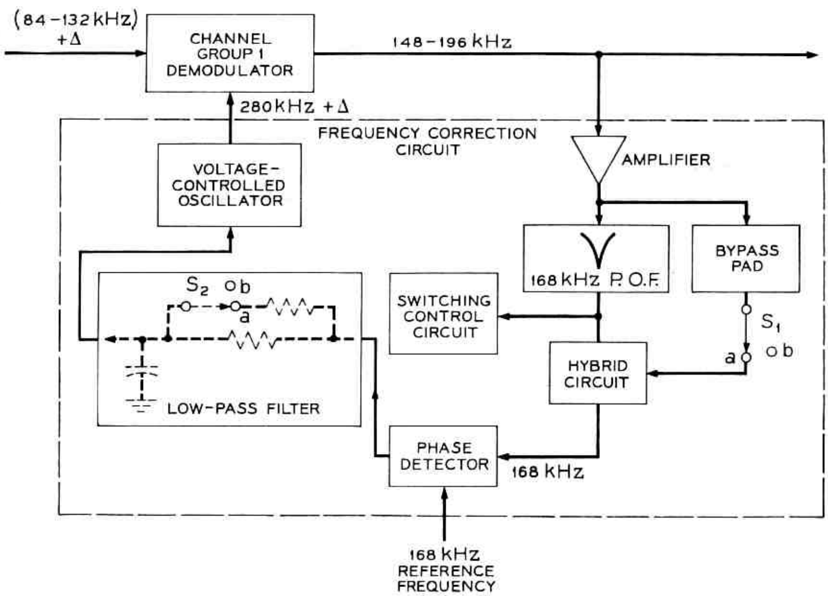

Fig. 14 — Block diagram of frequency correction circuit.

circuit to provide broad selectivity in the vicinity of 168 kHz (at the output of the channel group demodulator, one of the transmitted carriers is nominally at 168 kHz). Since the 168-kHz pick-off filter is only a few cycles wide, transmission is blocked through that path. The amplifier output signal is transmitted through the bypass pad, switch S_1, the hybrid circuit and into the carrier leg of the phase detector. The 168-kHz signal drive frequency for the phase detector is supplied from the precise N3 common carrier supply. The very low frequency output (nominally dc) is passed through the low-pass filter. Since switch S_2 is in the (a) position the low-pass filter is relatively wide band. The output provides the control signal for the voltage-controlled oscillator. This complete loop represents the wide-band loop which acts to bring the output of the voltage-controlled oscillator to a frequency of 280 kHz $+ \Delta$. As this point is approached, the 168-kHz carrier tone passes through the pick-off filter to the hybrid circuit. The switching control circuit senses this energy and after a short time delay switches S_1 and S_2 to the (b) positions thereby switching out the bypass pad and switching in the very narrow low-pass filter. The loop is now in the narrow-band mode and presents a high degree of discrimination to signals in the vicinity of the 168-kHz carrier. Therefore, if the incoming signal is shifted by an amount Δ, the voltage-controlled oscillator provides a clean channel group demodulating carrier which is also shifted by an amount Δ. Thus, the desired output signal has been corrected in frequency for the line shift.

The open loop characteristics of both the narrow and wideband loop are shown in Fig. 15. Except for an inherent 6 dB per octave slope in the rest of the loop, these characteristics are controlled by the two low-pass filter configurations. As noted on the curves the capture range for the wideband loop is 500 Hz and only 30 Hz for the narrow loop. The 500-Hz range for the wideband loop is required to allow for initial, temperature, and aging variations on the free-running oscillator in addition to the line shift error.

Since the outputs of the carrier supplies are derived by harmonic generation, the error at each carrier frequency will be different depending upon the error in the base generators and the particular harmonic that the carrier represents. Therefore, a residual error remains after frequency correction at all carrier frequencies except that of the reference tone. Based on a maximum base generator error of ± 7 parts per million, the worst case residual error is less than 0.6 cycle per second.

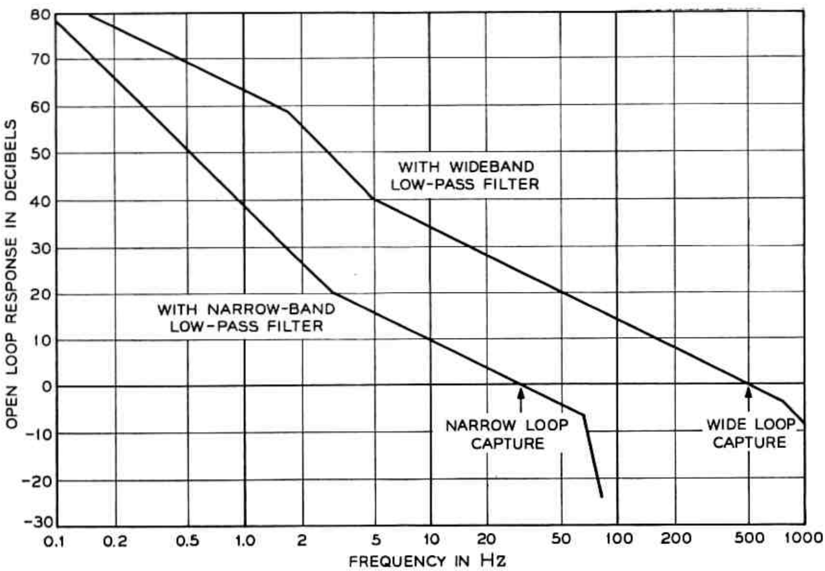

Fig. 15 — Open loop response of correction circuit.

VIII. GROUP TRANSMITTER CIRCUITS

The outputs of two channel group modulator circuits are combined in a hybrid coil network and applied to a group transmitter. This circuit may be required to transmit signals to the repeatered line in either the low group band (36–132 kHz) or the high group band (172–268 kHz). Therefore, two types of group transmitter units are available: a low group transmitter and a high group transmitter. The basic difference between the two types of group transmitter circuits is the inclusion of the modulator and associated circuitry in the high group transmitter. For this reason, only the high group transmitter shown in Fig. 16 will be discussed.

The single stage amplifier at the input provides approximately 11 dB of gain to the low-level low-group signal. The low-pass filter passes the 36 to 132-kHz band and rejects the unwanted modulation products generated in the channel group modulators. The slope equalizer allows the transmitted signal to be pre-equalized to partially compensate for the cable loss characteristic. Seven different equalizers are available for insertion in the circuit to provide a choice in slopes of 0, ±3, ±6, or ±9 dB for the channel 12 carrier of channel group 2 relative to the channel 1 carrier of channel group 1. The output of the equalizer is ap-

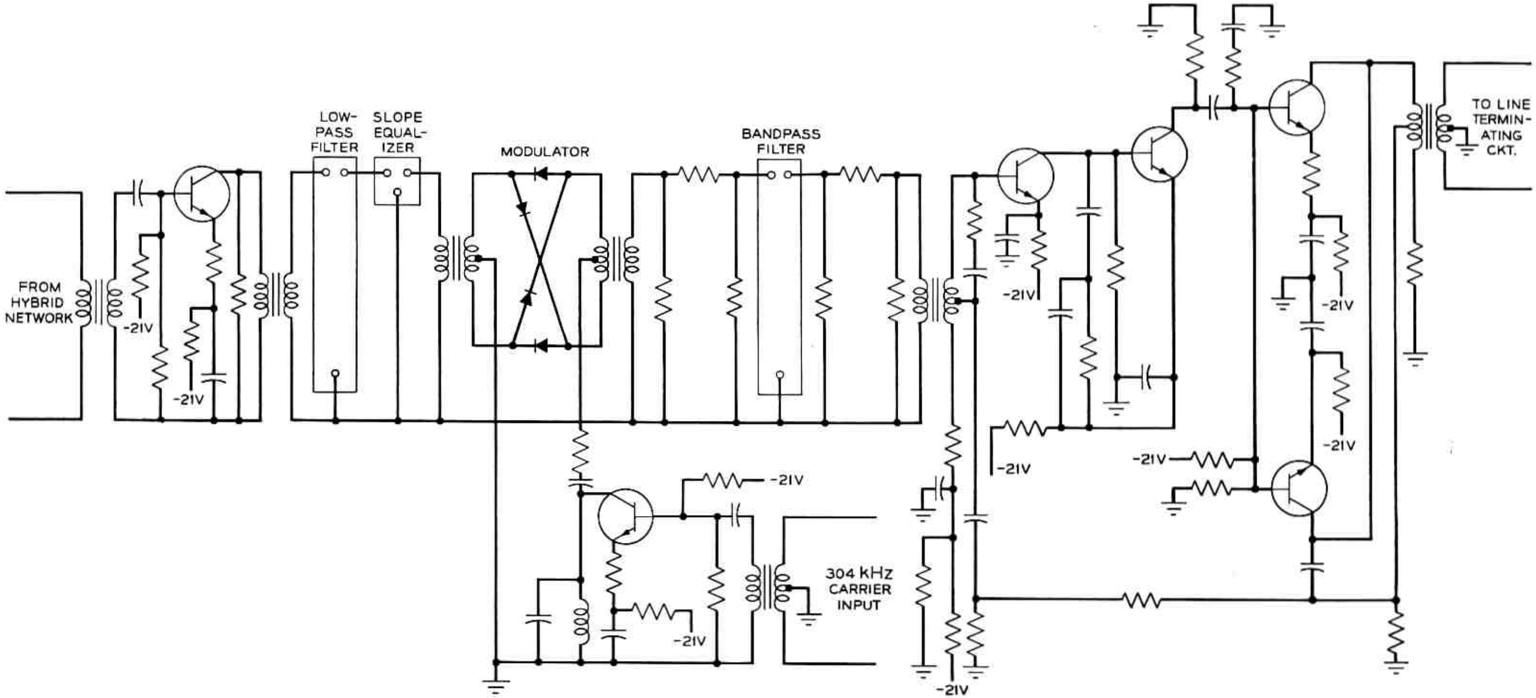

Fig. 16 — High group transmitter.

plied to the group modulator which consists of four diffused junction silicon diodes connected as a double-balanced ring modulator. The 304-kHz carrier for the group modulator is supplied from the common carrier supply and is amplified by a single stage buffer amplifier between the carrier supply and modulator. The bandpass filter following the group modulator transmits the high group band and rejects the unwanted modulation products from the modulator. The 3-stage amplifier provides approximately 60 dB of gain to the low level signal to obtain the proper levels for transmission over the carrier line. The third stage of the amplifier utilizes two transistors operating in parallel in order to handle the high output power levels with adequate intermodulation performance ($3F/F < -105$ dB for $F = 0$ dBm). As shown in Fig. 17, approximately 35 dB of negative feedback is employed both to improve the intermodulation performance and to reduce gain variations. At the high end cut-off better than 60 degrees of phase margin is achieved in addition to 17 dB of gain margin to insure the closed loop stability of the amplifier. The input and output impedances of the amplifier are precisely controlled by connecting the feedback circuit in the hybrid configuration. The hybrid connections also serve to stabilize the feedback independently of the amplifier terminations.

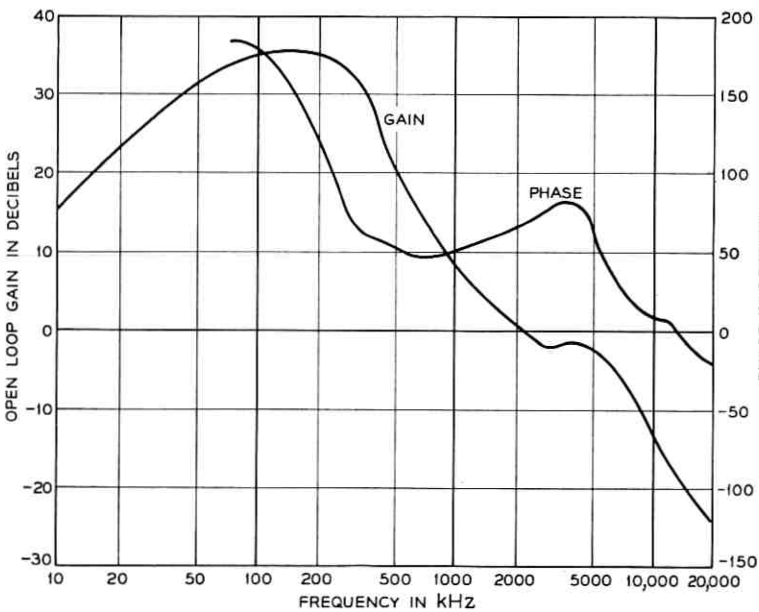

Fig. 17 — Open loop phase and gain response for hgt amplifier.

IX. GROUP RECEIVER CIRCUITS

The incoming signal to an N3 terminal may be in either the low group band (36–132 kHz) or the high group band (172–268 kHz). Like the group transmitters, it is necessary to provide two types of group receiver circuits: namely a low group receiver and a high group receiver. These two circuits are basically the same except for the inclusion of a demodulator and associated circuits in the high group receiver. Only the low group receiver shown in Fig. 18 will be discussed.

The low-pass filter at the input passes the low group band and rejects the unwanted out-of-band frequencies which may have been picked up due to crosstalk or other sources in the incoming line. As in the group transmitter circuit, slope equalizers are available for insertion in the circuit to compensate for the slope across the band due to the cable loss characteristic. In addition, the incremental slope adjustment circuit can be adjusted on a screw option basis to provide 0, +1, or −1 dB of slope in the gain characteristic across the low group band. The 3-stage amplifier following the slope equalizer provides approximately 53 dB of gain to the low level signals. A parallel output stage is employed to improve the intermodulation distortion performance. Shunt feedback is employed at the output of the amplifier while hybrid feedback is employed at the input for low noise performance and a good return loss. The shunt feedback connection at the output produces the low output impedance necessary for achieving isolation between the two channel group demodulators. The unusual shunt feedback connection on the output utilizes the transformer as an auto-transformer to provide the proper signal level to the thermistor. The resistor connected between the primary winding and ground is used for low frequency shaping. The open loop characteristics are similar to those shown for the group transmitter amplifier. However, due to the thermistor and slope adjustment circuit in the feedback path, wider variations occur in the gain and phase margins as these elements are selected or change value. The feedback circuit includes a thermistor for automatic gain control and an incremental slope adjustment circuit which is used as a fine adjustment to interpolate between the inserted equalizer values. The impedance of the thermistor is determined by the total output power. Thus, as the output level increases, the thermistor impedance decreases and the gain of the amplifier decreases tending to restore the output to its nominal level. This action reduces input level changes of ±8 dB to less than 1 dB at the output.

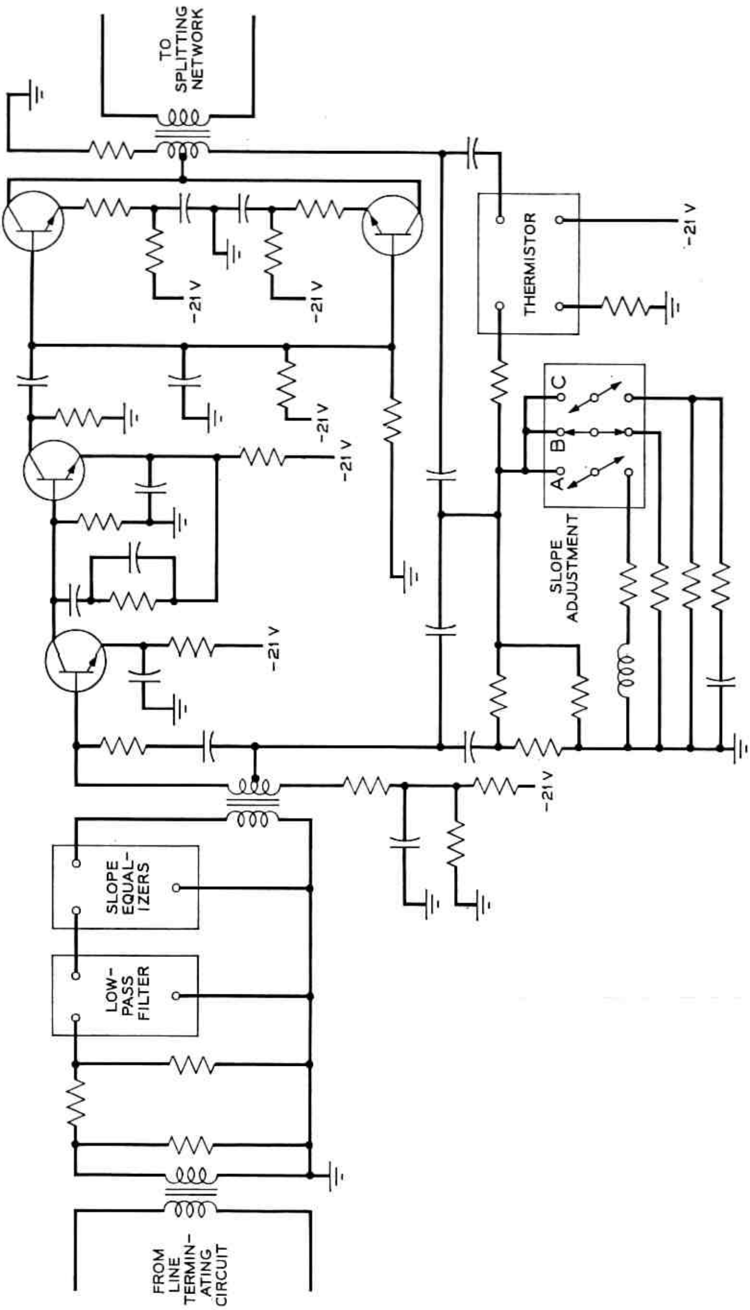

Fig. 18 — Low group receiver.

X. ALARM AND RESTORAL CIRCUIT

Separate carrier alarm circuits are provided for each channel group. Upon recognition of a received carrier failure, a carrier failure is forced at the far terminal and provision is made to release customers and apply a busy signal on the 12 channels. The circuit automatically monitors the transmission such that when the fault has been cleared the 12 channels at each end of the system are simultaneously restored to service.

The alarm and restoral circuit in conjunction with the external trunk release and make busy circuit automatically seizes control of the entire channel group in the event of a failure. When the total carrier power monitored at the output of the channel group demodulator circuit falls below the threshold level, the carrier office alarms are operated. The input to the channel group modulator is shorted for a period of 10 seconds to force a carrier failure in the associated alarm circuit at the far end of the system. By means of a ground signal, all 12 trunks are made idle. After a 10-second delay these trunks are made busy. The alarm circuits at this point are locked up and service cannot be restored until the automatic transmission tests are satisfactorily completed. A 2600-Hz tone is applied to the input of the channel modulator in the first test channel. The output of the channel demodulator is monitored for receipt of the 2600 Hz with an adequate signal-to-noise ratio. When this signal is satisfactorily received, the circuit in a similar fashion applies the tone to the input of the second test channel and monitors the output. When the 2600-Hz signal is satisfactorily received at the output of the second test channel, the signals to the make-busy circuit are removed and the system is automatically restored to service. An optional arrangement permits a single alarm and restoral circuit to control both channel groups in a terminal. In this case, however, all decisions are made on the transmission performance in only one of the two channel groups.

XI. N3 CARRIER-FREQUENCY SUPPLY

A novel feature of the N3 carrier system is the use of a carrier-frequency supply unit which furnishes all of the carriers needed for as many as 26 terminals. This common carrier supply provides the frequency stability and the amplitude control deemed necessary for a modern short-haul system and is economically competitive with other methods of carrier generation in all but the smallest installations.

Noteworthy features which enhance the operating flexibility of the new carrier supply include continuous monitoring to detect substandard performance, built-in automatic protection against service interruptions, convenient aids for easy maintenance, and alarms for all trouble conditions. The circuit design takes advantage of solid-state technology to minimize power consumption and space requirements. The N3 modulation plan[1] requires the generation of 16 different frequencies needed for modulation and demodulation. The channel modulators require 12 frequencies equally spaced at 4-kHz intervals and ranging from 148 kHz to 192 kHz. Two frequencies, 232 and 280 kHz, are required for the channel group modulators. One frequency, 304 kHz, is required for the group modulator. The one remaining frequency, 256 kHz, is needed for a modulator which will be part of the equipment for interconnecting the N3 and L systems. This modulator will shift the frequencies in the N3 channel group band, 148 to 196 kHz, to the A-type channel bank range (L multiple group band), 60 to 108 kHz. Thus, the N3 and L systems can be interconnected at channel bank frequencies rather than at voice frequencies.

The generation of many carrier frequencies, accurately and economically, is a basic need for any system and is a crucial requirement for a short-haul system. A study of the frequency stability attainable with crystal-controlled oscillators, similar to those used in the existing short-haul carrier systems, revealed a need for smaller frequency differences from terminal to terminal. Consideration of system performance objectives led to the decision to provide a high-grade 4-kHz oscillator at each location. By allocating all carrier frequencies to exact multiples of 4 kHz, the carriers can be generated as harmonics of a 4-kHz base frequency.

Synchronization of the carrier oscillators at the two terminals is hindered by the random frequency shifts introduced by the local oscillators used for transposing and inverting the channel groups at each repeater, i.e., frequency "frogging." Since a change in the repeaters is impractical, the only alternative is to correct the received carrier frequencies at each terminal. Frequency correction can be achieved rather easily if all of the carrier frequencies are harmonics derived from a single primary frequency source. If the carrier frequencies at the transmitting terminal and at the receiving terminal are obtained from similar harmonic generation processes, the terminal-to-terminal frequency errors can be made dependent only upon the two relatively precise base frequency oscillators at the transmitting and receiving terminals.

The foregoing considerations are the basis for the specification of

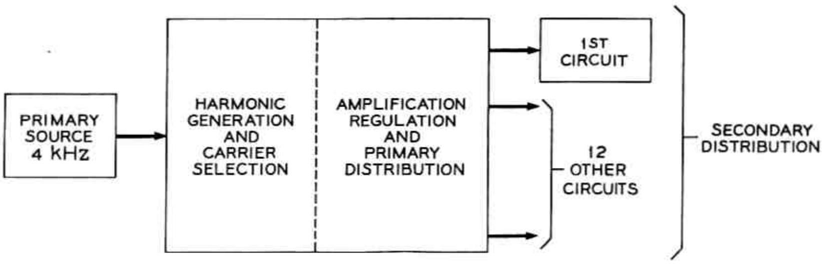

Fig. 19 — Carrier-supply — general plan.

requirements for the N3 carrier-frequency supply. The block diagram of Fig. 19 shows the functional plan. A primary frequency source operating at 4 kHz serves as an input to the carrier generator circuits. The primary frequency supply for the L Multiplex is a preferred source, when available, but a stable crystal-controlled oscillator has been designed for the N3 carrier-frequency supply to meet a frequency stability objective of ± 7 parts per million over an ambient temperature range from 32° to 120°F and including aging for a 12-month period.

All of the carriers are supplied to the terminal bays from the primary distribution panel in the carrier frequency supply bay. Each of the arrows shown in Fig. 19 represents a multiconductor cable which connects one secondary circuit to the common carrier supply and transmits 16 carrier frequencies. Each secondary circuit is located conveniently in a carrier terminal bay and is capable of supplying all carriers for two terminals.

The carrier voltages at the output of the primary distribution panel are relatively pure sine waves. The second and third harmonic content is at least 60 dB below the fundamental. All unwanted carrier frequencies are at least 58 dB below the wanted frequency at each output. Each carrier is regulated in amplitude to be within ± 0.5 dB of the nominal voltage. The 12 channel carriers are distributed at a level of +11 dBm into each 115-ohm circuit. The other carriers are furnished at a level of +8 dBm. At the input of each secondary distribution circuit, the amplitudes are adjusted to equalize the losses in different lengths of cable and to be within ± 0.5 dB of the nominal voltage. The six transmitted carriers which are used for regulation are maintained within limits of ± 0.1 dB. Although only 16 different carrier frequencies are distributed, 30 outputs are needed to furnish all of the carriers required for each terminal.

Provision of a common carrier supply for as many as 624 channels increases the importance of continuity of service. To provide the high degree of reliability required, it is essential that it be possible to duplicate all active components. A "failure" should cause an automatic transfer from the regular to the standby equipment. The automatic transfer to standby equipment should not be made unless the standby equipment is operating satisfactorily within the specified limits. Since such redundancy is costly, and in the smaller installations may be unwarranted, the addition of the standby equipment is made optional. Provision is also made for the addition of standby units when needed without requiring changes in wiring.

XII. GENERAL PLAN FOR CARRIER-FREQUENCY SUPPLY

Analysis of the requirements given in the preceding section leads to a general plan for the carrier-frequency supply. The important relationships can be placed in evidence by considering three functions separately. (*i*) Generation: Sixteen carrier frequencies must be produced as harmonics of a single 4-kHz input voltage from a primary source. (*ii*) Distribution: A large number of independent output voltages (1482) must be distributed to 26 terminals and each voltage must have a specified magnitude and frequency. (*iii*) Supervision: A comprehensive monitoring and automatic switching and alarm system must be furnished to assure continuity of service and to provide a visual status display for supervision.

12.1 *Primary Frequency Supply*

The primary frequency source for the carrier supply comprises a 4-kHz oscillator and an amplifier that provides sufficient excitation current for the harmonic generator. The essential components of the circuit are shown in Fig. 20. The dotted lines show the optional standby equipment that will provide protection against failure of a fully-equipped system. When the 4-kHz signal is available from the L Multiplex carrier supply, it is fed directly to the inputs of the two amplifiers. A built-in crystal-controlled oscillator is furnished whenever the L Multiplex carrier supply is not available and an internal frequency source for a self-contained system is required. In either case, both amplifiers are fully energized. The working amplifier drives the harmonic generator, and the standby amplifier delivers full output to a resistive load at all times. Each amplifier is equipped with monitoring facilities which determine whether the output current is within specified limits.

The amplifier shown in the block diagram of Fig. 20 is assembled

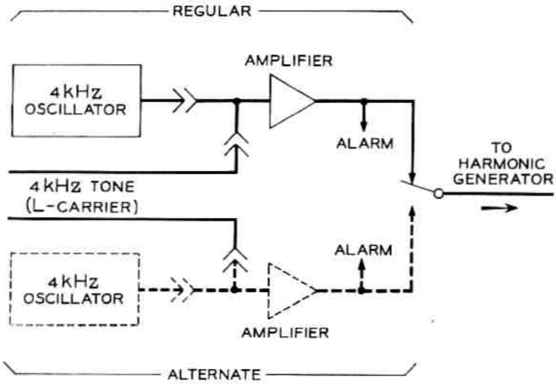

Fig. 20 — Primary 4-kHz source.

as part of the plug-in unit designated as the 4-kHz generator. The primary function of the circuit is to provide current to excite the harmonic generator. A second important function is to provide information to the switching and alarm system whenever the 4-kHz output current is not adequate to excite the harmonic generator properly. A relay initiates office alarms and an automatic switching transfer whenever the output is not within prescribed limits.

The 4-kHz amplifier circuit is shown in Fig. 21. It was designed to

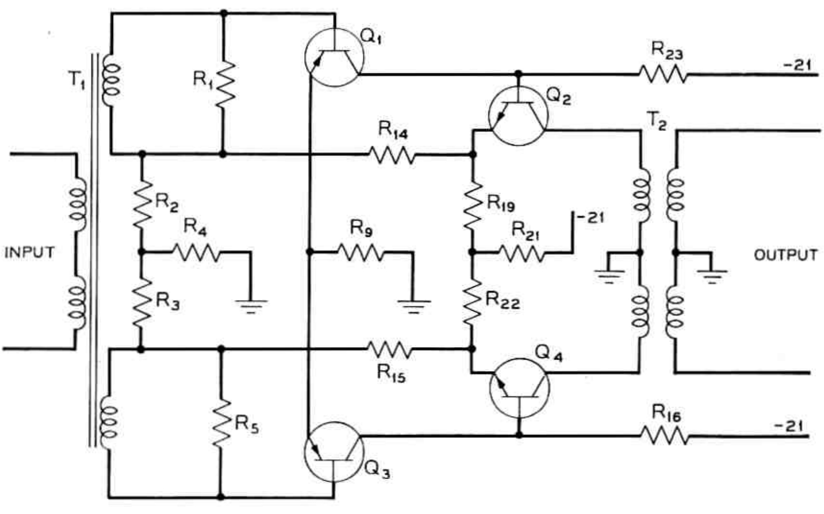

Fig. 21 — 4-kHz amplifier.

deliver sufficient output current to excite the harmonic generator when the input current is supplied by either the L carrier multiplex or the self-contained local oscillator. Both sources are capable of delivering 0 to 2 dBm to a 135-ohm load. However, the local oscillator must energize an oven alarm circuit which requires about half of the available power. For this reason, the amplifier was designed to present a 270-ohm impedance to the source and a 270-ohm shunt resistance is added whenever the power is supplied by the L-carrier multiplex. Thus, the input current to the amplifier is 1.5 ± 0.2 milliamperes. Since the current gain of the amplifier is designed to be 24 dB, the output current delivered to the harmonic generator is between 22 and 28 milliamperes for the range of inputs described above. The harmonic generator can be excited with smaller currents but this range is preferred.

The harmonic generator circuit includes a saturable magnetic core inductor, and the input impedance changes depending upon the magnitude of the input current. Since the harmonic generation process is controlled by the magnitude of the magnetizing force, a sinusoidal current waveform is preferred. Therefore, the negative feedback amplifier uses current feedback from the output mesh to maintain an output current amplitude that is relatively insensitive to changes in the impedance of the load. The ac voltage between the emitters of transistors Q_2 and Q_4 is proportional to the current through the output transformer. A fraction of this voltage is fed back to the input mesh. Since the feedback is proportional to the output current, the amplifier will approximate a constant-current source for exciting the harmonic generator. Thus, a sinusoidal input produces a sinusoidal output current even though the load impedance is nonlinear.

12.2 *Harmonic Production and Selection*

The harmonic generator and the filters which select the carriers for the N3 carrier supply are shown in Fig. 22. The harmonic generator circuit employs a saturable core inductor and produces an output waveshape that is very rich in odd harmonics of 4 kHz. Although the even order components on the spectrum are subject to some variation depending upon the symmetry of the hysteresis loop, they are at least 25 dB below the odd order components at the input to the filters. To produce even harmonics of 4 kHz, a full-wave rectifier is connected across the odd harmonic output. Fig. 22 shows the circuit that produces the harmonics and the bandpass filters that select the wanted carrier frequencies from the harmonic spectrum. By arranging the filters in two

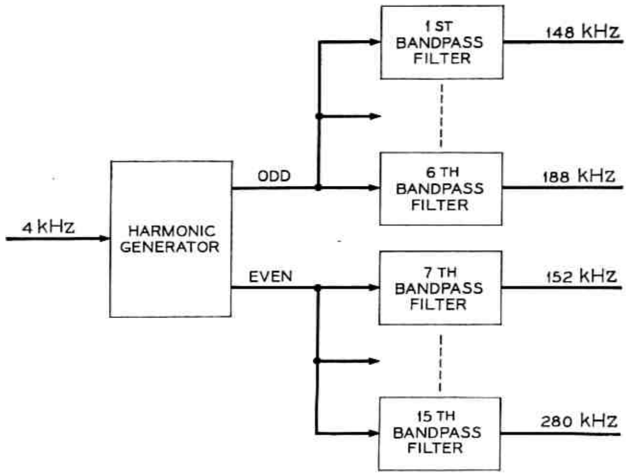

Fig. 22 — Harmonic generation and selection.

groups, one connected to the odd harmonic circuit and the other connected to the rectifier circuit, the unwanted components at the input to each filter are at least 8 kHz away from the wanted component, and the filter requirements can be relaxed. Six filters, with their inputs in parallel, are connected to the odd harmonic output terminals. Similarly, nine filters are connected to the even harmonic output terminals. If the diode bridge used for rectification is well balanced, the odd harmonics of 4 kHz are at least 25 dB below the even order harmonics at the input to the filters. The bandpass filters use crystal units to obtain narrow passbands with adequate out-of-band attenuation.

The circuit used for the simultaneous generation of a large number of harmonics of approximately equal amplitude is shown in Fig. 23. The essential circuit element is the nonlinear inductor, L_4, a coil which is operated with sufficient magnetizing force to drive its magnetic core material well into the saturated region. Thus, the inductance of the coil is large when the current through it is small, and its inductance is small when the current through it is large. The main performance features of the circuit can be reproduced by a crude model which attributes to the inductor, L_4, a sort of switching property. The primary current is constrained to be sinusoidal by the 4-kHz feedback amplifier. The impedance of the coil is high near the zero crossings of the current wave, and the capacitors, C_4 and C_6 are charged slowly. The impedance is low throughout the peak of the current

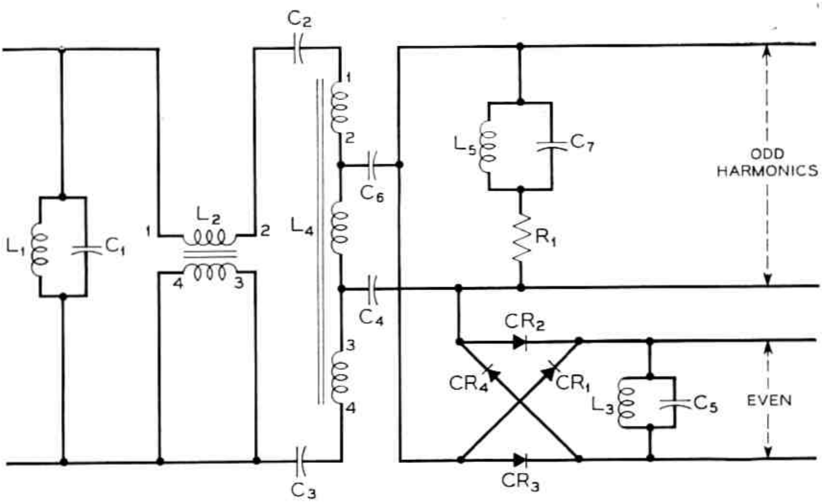

Fig. 23 — Harmonic generator circuit.

wave, and the capacitors discharge rapidly through the resistance of the load and the small inductance of L_4.

The waveforms shown in Fig. 24 have been drawn to illustrate the sequence of events during a complete cycle of the fundamental input wave. A simplified version of the harmonic generator circuit shown in Fig. 23 has been drawn to facilitate the following discussion. The output current pulse that is characteristic of this type of harmonic generator is actually more sharply peaked than the sketch shows. It is the narrowness of the discharge pulse that provides the principal contribution to the higher harmonics which are needed; and the circuit parameters are adjusted to maintain the harmonic distribution as uniform as possible over the frequency range of interest.

Within the interval (t_0, t_1), the amplitude of the sinusoidal current I_1 exceeds the threshold $+I_0$. The nonlinear inductor L_4 is in the saturated state where its inductance is low and the voltage drop across it is correspondingly small. The current I_2 which charges capacitor C_4 is very small. Within the interval (t_1, t_2), the amplitude of the current I_L is not sufficient to cause saturation. The inductance of L_4 increases suddenly at the threshold and the voltage drop across it increases. The current I_2 which charges capacitor C_4 increases. Charging continues until the current through L_4 is sufficient to cause it to switch to the saturated state at time t_2. The capacitor discharges

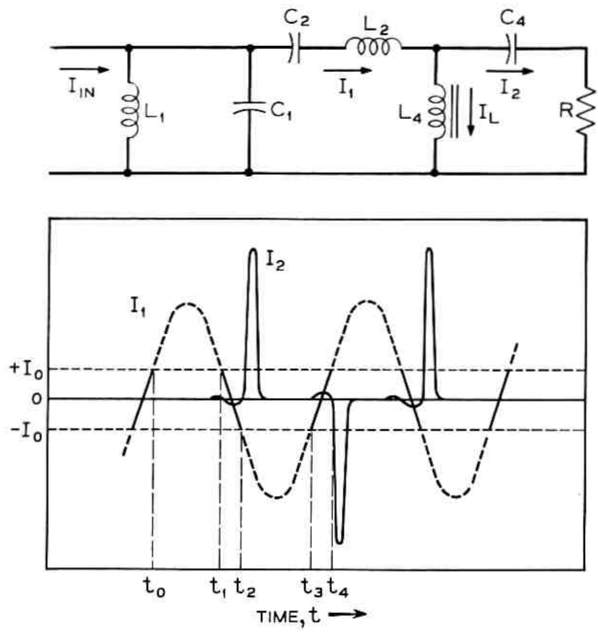

Fig. 24 — Typical pulse waveforms.

through the load resistor. The resistance, capacitance, and saturated inductance effectively in the circuit are adjusted to permit the current to rise to a high maximum, to damp the pulse, and to shorten the pulse duration to the point at which the highest harmonic required reaches the desired amplitude. The pulse dies away before the end of the cycle at time t_3. At that time, the currents and voltages are the same, except for reversals of sign, as those at the start. So, the current wave consists of an alternating succession of these positive and negative pulses. Thus, only odd-order harmonics are generated when the core of the nonlinear coil is unpolarized, as is the case here.

12.3 *Amplification and Regulation of Carriers*

The harmonic generation and selection process yields tones of the required purity but the available power is subject to moderate variation. Since the output of each of the crystal filters is low, an amplifier is required to provide sufficient power for 26 terminals. Therefore, each amplifier has been designed to provide regulation as well as power gain. Fifteen amplifiers receive inputs from the 15 bandpass filters. The one remaining amplifier receives its input from a frequency doubler; thus, an input of 152 kHz produces an output of 304 kHz. The sixteen amplifiers are all alike and each is equipped with monitoring facilities which determine whether the output current is within specified limits. A relay initiates office alarms whenever the output is not within proper limits.

Two independent amplifiers and a common monitoring circuit comprise a plug-in assembly which has been designated as the dual amplifier unit. Two dual amplifier assemblies, a working unit and a standby unit, are connected as shown in Fig. 25. The "failure" of either amplifier in the regular dual amplifier unit causes an automatic transfer to the alternate unit.

The circuit for one of the amplifiers on a dual amplifier unit is shown in Fig. 26. The nominal input to an amplifier is -5 dBm, but it may be as low as -14 dBm or as high as 0 dBm. For any input within this range, the power output is required to be between $+22.5$ and $+23.5$ dBm. Although the amplifier must provide sufficient gain to satisfy the power requirements of the carrier system, its most important performance characteristic is the ability to regulate automatically the amplitude of the output voltage.

The mechanism by which the regulation is achieved is an interesting feature of this circuit. The level of the input signal is raised by the gain of transistor Q_1 to provide a strong current drive for the "push-pull" output stage. During a part of each cycle when the signal is close to a zero crossing, the output stage acts as a linear amplifier and produces a corresponding voltage across the load. However, this con-

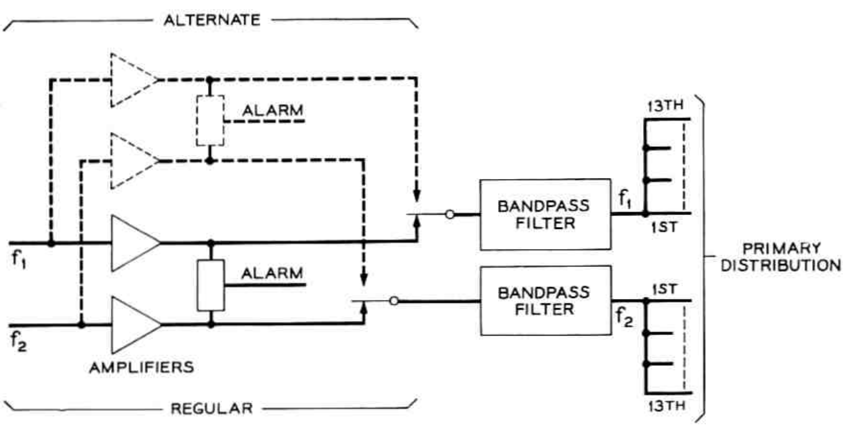

Fig. 25 — Amplification and distribution.

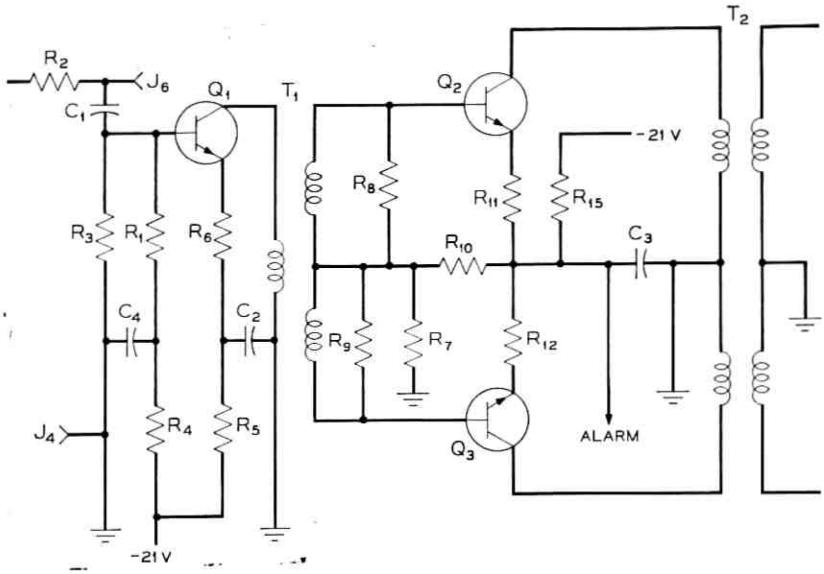

Fig. 26 — Regulating amplifier.

dition lasts for a small fraction of the cycle because bias and load have been chosen to force one of the transistors into saturation and the other into cut-off. During the subsequent interval, almost all of the voltage across capacitor C_3 appears across one-half of the primary winding of transformer T_2. The energy stored in C_3 is adequate to maintain a nearly constant voltage across the load during this interval. The process is repeated during the next half-cycle with the voltage being applied across the other half of the primary winding. The net effect is to produce across the load an alternating voltage having a trapezoidal waveform. The magnitude of this voltage is determined almost entirely by the voltage on capacitor C_3. Resistors R_{11} and R_{12} aid in producing odd symmetry in this waveshape; and hence, reduce the even-order harmonics in the output. The magnitudes of the odd-order harmonics will, of course, be substantial.

This mode of highly nonlinear operation produces a distorted output waveform; but a relatively simple filter will separate the fundamental component from the harmonics when the input signal is a single frequency sinusoid. Measurement shows that the amplitude of the fundamental component is almost constant for a wide range of input signal magnitudes.

During the operation of the circuit as a regulating amplifier, the transistors switch current pulses periodically from one half of the primary winding of transformer T_2 to the other. Alternating current flows in the output circuit of the transformer, and a voltage builds up across capacitor C_3. The duration of the current pulse depends upon the amplitude of the input signal, but the peak magnitude of the pulse is almost proportional to the voltage across capacitor C_3. Regulation is achieved by allowing this voltage to increase when the pulse duration is shortened; i.e., when the input signal decreases. This is accomplished by supplying all the current through resistor R_{15} which is connected to the -21 volt supply. The constant component of the voltage across C_3 depends upon the area of the current pulses, and this voltage must equal the supply voltage less the voltage drop across resistor R_{15}. This resistor, together with trimming resistors as required by a factory adjustment, provides the means for adjusting the output voltage to the required magnitude.

The output of the amplifier is delivered to a bandpass filter which is part of the primary distribution circuit. The unwanted harmonic voltages are suppressed, leaving a relatively pure sine wave of the wanted frequency. By placing the filter on the primary distribution panel, the plug-in amplifier assemblies can be made alike. They can be used interchangeably in all positions requiring the dual amplifier. Fig. 27 shows typical measurements of the variation in output power

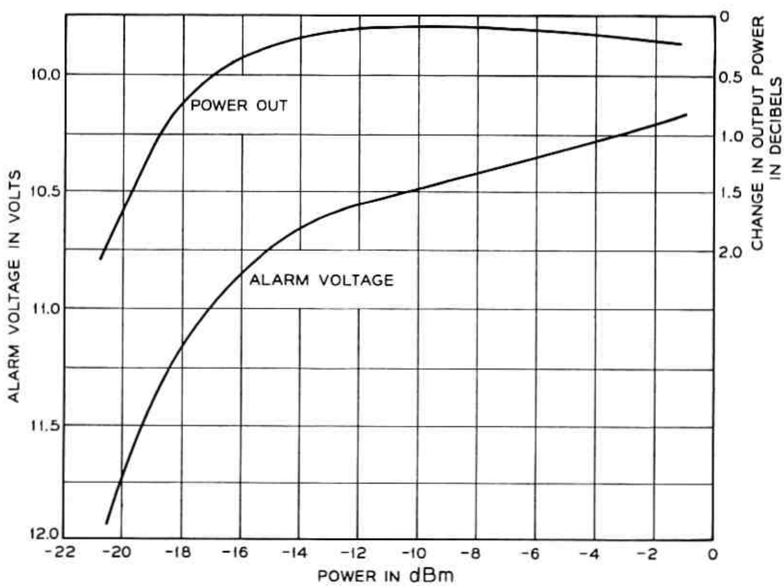

Fig. 27 — Regulating amplifier output power as a function of input signal.

as the input signal is varied over a wide range. Since the measurements were made at the output of the primary distribution system, the filter has selected the fundamental and suppressed the harmonics. The frequency is the same as the input signal, but the variations in amplitude have been drastically reduced.

Fig. 27 also shows the variation in the voltage across capacitor C_3. The magnitude of this voltage is directly related to the output voltage; and thus, it is a convenient means for monitoring the amplifier output. The curve has been labeled "Alarm Voltage" because this voltage is used to initiate alarms when the output power is not within specified limits.

The two regulating amplifiers in the dual amplifier unit share the common monitoring circuit shown in Fig. 28. The diodes CR_1, CR_2, CR_3 and CR_4 together with the associated resistors form a logic circuit. If the voltage on either the ALM IN A or the ALM IN B leads increases or decreases sufficiently, one diode conducts and applies a signal to the summing amplifier. The output of the summing amplifier which causes the trigger circuit to operate the relay K_1 is taken from the collector of Q_5. Either an increase on the Q_4 base or a decrease on the Q_5 base causes the collector voltage of Q_5 to decrease. Thus, the circuit permits establishing both upper and lower limits; and the relay operates whenever the signal is not between the limits. Transistors Q_9 and Q_{10} are connected to form a Schmitt trigger circuit. Regeneration occurs when the gain of the positive feedback loop

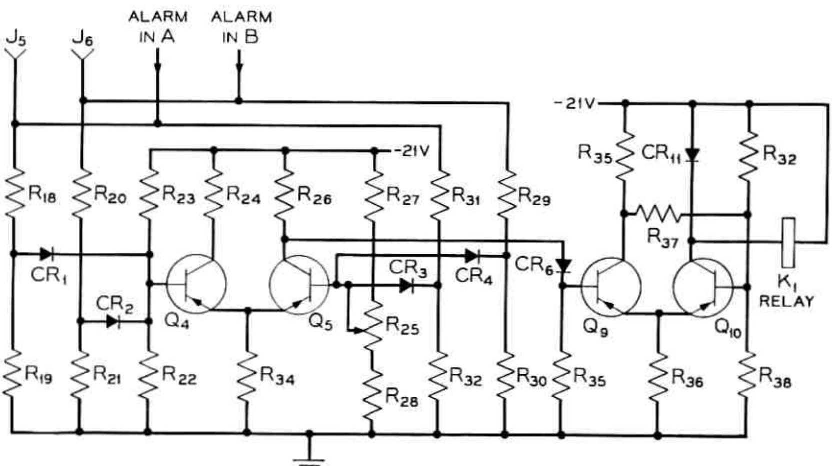

Fig. 28 — Monitoring circuit.

exceeds unity. The switching process proceeds rapidly when the threshold voltage is crossed.

XIII. PRIMARY DISTRIBUTION CIRCUIT

The distribution of carrier-frequency voltages to as many as 26 carrier terminals requires a large number of cable pairs. The number increases rapidly with the number of carrier terminals; and to achieve economy and convenience, the distribution is accomplished in two stages. A primary distribution unit which is part of the carrier supply assembly may be connected to as many as 13 secondary distribution units. One secondary distribution panel is furnished for every two terminals, and it supplies all of the carrier voltages needed. The 16 carrier frequencies are fed to each secondary distribution panel on 19 pairs in an interconnecting cable. Two pairs are used for each of the two channel group carriers and for the group carrier. One pair is used for each of the 12 channel carriers and for the N3 to L translation carrier. An adjustable pad is provided at each of the 19 inputs to the secondary distribution unit. These pads provide a means for compensating for the losses in the interconnecting cable which may vary in length from a few feet to a maximum of 700 feet.

The primary distribution panel provides means for connecting 13 secondary distribution units to the carrier supply. The circuit includes 16 bandpass filters, one being used for each carrier frequency. Each filter suppresses the unwanted harmonics generated in the regulating amplifier and delivers a sinusoidal wave form having the wanted carrier frequency to a multiplicity of resistance loads. The resistance loads are connected through capacitors that are part of the filter structure. A simplified schematic circuit for the primary distribution of one of the 16 carrier frequencies is shown in Fig. 29.

The principal function of each filter is to suppress harmonics, and a low-pass filter would be adequate for this purpose. However, the net-

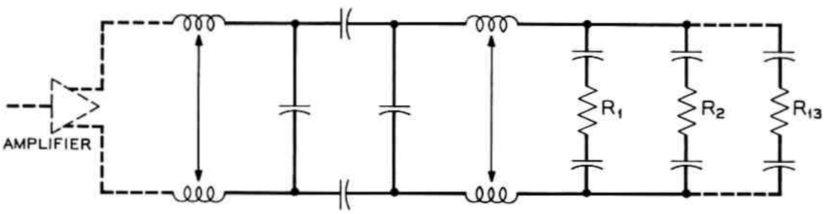

Fig. 29 — Primary distribution circuit.

work used is an impedance transforming bandpass filter designed for insertion between a source impedance, R, and an impedance, R/N, which represents N loads connected in parallel. The output loads are independent circuits, and it is important to provide isolation between them so that one will not react upon another. This has been accomplished by placing a capacitor in series with each load. By designing the filter to include the capacitor, sufficient isolation is achieved without loss of power. The filter is adjusted to present a resistive termination of 115 ohms to the limiting amplifier when each output load is 115 ohms. All of the outputs are terminated in either a cable pair having a 115-ohm load at the far end or with a 115-ohm resistor which is furnished with the unit. These resistors are removed by cutting leads when a cable is connected.

XIV. SECONDARY DISTRIBUTION CIRCUIT

Although only 16 different carrier frequencies are needed by each carrier terminal, many different voltages are required for the circuits in each terminal. A fully-loaded carrier frequency supply and the 26 carrier terminals that it serves would require 1482 interconnecting pairs. By locating a secondary distribution unit in each terminal bay the lengths of the interconnecting cable pairs are minimized. Arranging the distribution system in two stages is economical and also provides flexibility in locating the carrier terminals.

Each secondary distribution unit receives inputs from 19 pairs which connect it to the primary distribution panel of the common carrier supply. Two N3 terminals are connected to a secondary distribution panel with a total of 114 pairs. Fig. 30 shows the plan for the secondary distribution circuits.

As shown in Fig. 30, each of the 19 incoming cable pairs is connected to an attenuator. These pads are adjustable in 0.5 dB steps and are used to compensate for variations in losses in the cable between the primary distribution panel and the secondary distribution panel. The output distribution networks are designed to provide isolation between the outputs. Six carrier regulator circuits are used to control accurately the amplitudes of the six transmitted carrier frequencies. These regulators are similar to the output stage of the regulating amplifier shown in Fig. 26.

XV. SWITCHING AND ALARM CIRCUIT

As the number of carrier terminals connected to a common carrier supply increases, it becomes important to minimize the possibility of service interruptions. The more vulnerable components of the N3

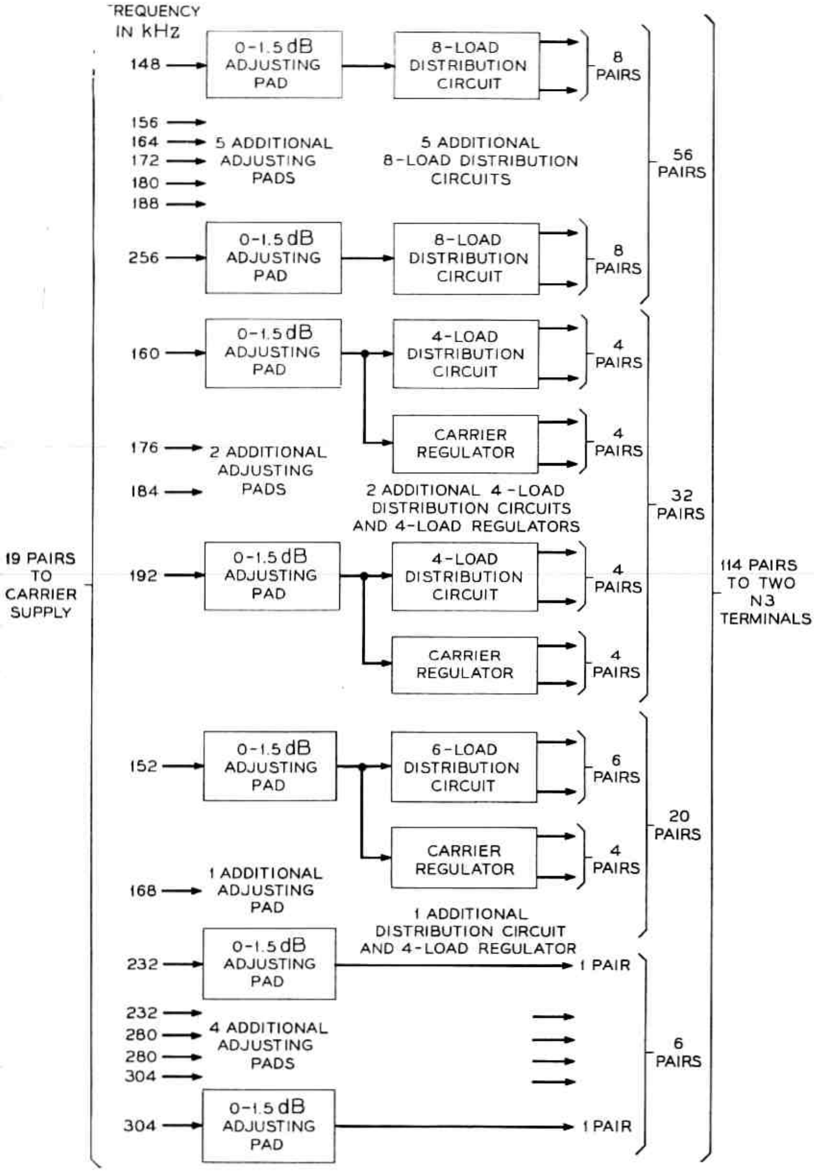

Fig. 30 — Secondary carrier distribution circuit.

carrier supply have been assembled as plug-in units that can be replaced easily. Also, the shelves on the bay frame have been equipped with pairs of receptacles wired so that plug-in components can be used in duplicate pairs to provide standby protection.

If duplicate units are provided, both the regular and the alternate units are maintained in an operating condition and their outputs are monitored at all times. Automatic detection and switching facilities select from each pair and connect into the system one unit whose output is within specified limits. Manual switches for producing a transfer are provided as an aid in maintenance. A trouble indicating lamp near each unit is lighted if the output is not within prescribed limits. Simultaneously, the appropriate major or minor alarm is initiated.

The principal function of the automatic switching circuit is to maintain continuity of service by transferring the load to a standby unit whenever a regular unit becomes defective. Another important function is to aid maintenance and supervision of the carrier supply by providing means for making a manual transfer from one unit to the other. The schematic diagram for the configuration of the contact network in the control path for a typical transfer relay is shown in Fig. 31.

The input conditions that initiate and control the action of the transfer relay are derived from the relays in the monitoring circuits on the regular unit (designated R), and the alternate unit (designated A), and a manual key on the switching and alarm panel (designated S). The schematic diagram represents normally open contacts by crosses and normally closed contact by bars. The R and A relays are in the normal released state when the units are within limits. As shown in the diagram, the S switch is selecting the regular unit, both the regular and the alternate units are within limits, the transfer relay is released, and the load is connected to the regular unit. Current

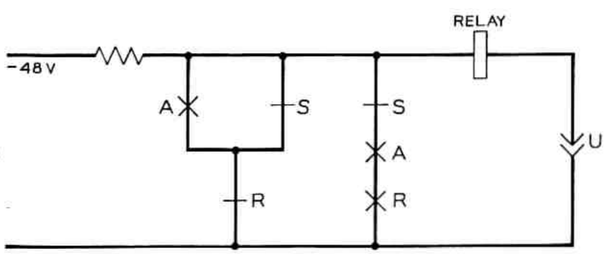

Fig. 31 — Relay control circuit.

flows through the closed R and S contacts in the left-hand path and the relay is prevented from operating. If the R unit fails, this shunt path is opened, and the relay operates to transfer the load to the A unit. If the S switch is turned to the alternate position, the shunt path is opened and a manual transfer is completed when the transfer relay operates.

The contact designated U is a link through the connector to the plug-in assembly. It is closed when the alternate unit is inserted. This interlock circuit inhibits either automatic or manual transfers whenever the carrier supply is not equipped with an alternate unit.

The right-hand path becomes effective whenever both units are out of limits. The key can be used to select either unit. This feature is desirable since the system may still be usable (with impaired performance) if one of the units is not too far out of limits.

Both the major and the minor alarm systems include a relay operated by a flip-flop circuit and means for generating a driving pulse whenever a monitoring relay operates in any of the protected units. After the alarm circuit has been operated by a pulse, the flip-flop circuit can be reset by using the reset button. Thus, a second pulse will produce a second alarm even though the first trouble has not been cleared.

The minor alarm is activated whenever either an R or an A relay operates on one of the plug-in units. The major alarm is activated whenever both the R and the A relays operate on any pair of plug-in units. Thus, a minor alarm means that at least one unit is working without standby protection. A major alarm means that at least one of the carrier frequencies is out of limits. When an alarm occurs, the lamps associated with the separate units indicate which units are out of limits. Replacement of a defective unit restores the system to the normal condition.

XVI. PRIMARY SOURCE OF POWER

The primary source of power for the carrier supply is the -48 volt plant battery. Since the types of transistors selected for the amplifiers cannot be operated efficiently with supply voltages much greater than 20 volts, a dc-to-dc converter is used to provide an efficient internal -21 volt source. Regulation and voltage transformation are accomplished by using a transistor as a switch. A control circuit causes a transistor to conduct and to be cut off periodically at a nominal switching rate of 10 kiloHertz per second. The output voltage is determined by the fraction of the total cycle during which the transistor is conducting. A large smoothing capacitor is charged by the current pulses; and the dc output voltage is proportional to the area of the pulses. The regulator circuit compares a fraction of the output voltage with a reference voltage provided by an avalanche diode. The comparison circuit changes the current pulse duration so that the desired voltage is maintained. The voltage is adjusted initially and is maintained thereafter within the range -21.0 ± 0.1 volts. Since the carrier voltages delivered to the primary distribution system are proportional to the voltage of this -21 volt supply, precise regulation is required. Both high-voltage and low-voltage alarms are provided. This is the same power supply used in the N3 terminal.

XVII. SUMMARY

A new 24-channel single-sideband terminal has been developed to supersede the existing ON terminal. The N3 terminal employs such new circuits as the common carrier supply and the frequency correction circuits to achieve its excellent performance. Full use has been made of the advances in the solid-state art, crystal filter design techniques, and high-performance miniature transformers. The N3 system performs significantly better than ON while being easier to install and maintain. It has been designed with a potential for full flexibility with the long-haul systems.

The first N3 terminals were placed into operation between Texarkana and Dallas, Texas on December 6, 1964.

XVIII. ACKNOWLEDGMENT

The circuits described here are the result of the combined efforts of many members of the Bell Telephone Laboratories. The authors wish to thank Messrs. F. H. Blecher, A. J. Grossman, D. D. Sagaser, and O. L. Williams for their direction and guidance.

REFERENCES

1. Bleisch, G. W. and Irby, C. W., The N3 Carrier System: Objectives and Transmission Features, B.S.T.J., This Issue, pp. 767–799.
2. Lundry, W. R. and Willey, L. F., The N2 Carrier Terminal — Circuit Design, B.S.T.J., *44*, May–June, 1965, pp. 761–785.
3. Giguere, W. J., Bowie, L. A., and Watters, J. A., Equipment Design of N3 Carrier Terminals and Carrier Supply, B.S.T.J., This Issue, pp. 845–894.
4. Boyd, R. C. and Herr, F. J., The N2 Carrier Terminal — Objectives and Analysis, B.S.T.J., *44*, May–June, 1965, pp. 731–759.
5. Boyd, R. C., Redesign of Short Haul Carrier Systems, Bell Laboratories Record, *41*, July–August, 1963, pp. 275–279.
6. Caruthers, R. S., The Type N1 Carrier Telephone System: Objectives and Transmission Features, B.S.T.J., *30*, Jan., 1951, pp. 1–32.

Equipment Design of N3 Carrier Terminals and Carrier Supply

By W. J. GIGUERE, L. A. BOWIE and J. A. WATTERS

(Manuscript received August 23, 1965)

This paper discusses the equipment aspects of the 24-channel N3 carrier terminals and associated common carrier supply. Arrangements of the carrier equipment, in shop-wired and tested packages with associated equipment, are described. The common carrier supply equipment is reviewed. Particular attention was given to those equipment features which have to do with reliability, economy, and installation and operating convenience.

A variety of package arrangements are described and the numerous advantages resulting from the "shop-wired package" concept are explored. The equipment features of the plug-in units which are components of both carrier terminal and carrier supply are examined.

I. INTRODUCTION

The N3 carrier terminal is a 24-channel, single-sideband, transistorized short-haul carrier terminal designed for use with carrier lines using N1, N1A, or N2 repeaters. It may be connected to pairs in the same cables with N1, N2, ON1, and ON2 systems. It is the successor to the ON2 carrier terminal and is designed to meet the latest performance requirements for two-way direct distance dialing or voice band data channels. The new N3 terminals provide important transmission improvements, as compared to ON2 terminals, which are discussed in companion papers.[1,2] For reasons of economy and improved performance, these new terminals are arranged in shop-wired packages together with functionally related signaling, trunk processing, carrier supply, and channel patching jack equipment.

This paper will discuss equipment design features of both the carrier terminal and the associated common carrier supply as they affect transmission performance, economy of manufacture, engineering, installation, maintenance, operating convenience, and system reliability.

II. GENERAL DESCRIPTION

Each 24-channel terminal includes two 12-channel groups rather than the six 4-channel groups employed in the ON2 terminal. This reduces the cost per channel through a reduction in the amount and cost of common group equipment. Additional advantages in having two 12-channel groups will be discussed in greater detail in companion papers appearing elsewhere in this issue.[1,2]

The terminals are arranged to include independent 12-channel group alarm and restoral units and associated trunk release and make busy units. This equipment functions to provide an alarm in the event of system failure, to properly process the associated trunks when failure occurs, and to restore the trunks to service automatically upon system restoral.

The N3 terminals obtain the various group, channel group, and channel carrier frequencies from a common carrier supply which is designed to provide the necessary frequencies for up to twenty-six N3 terminals. The carrier supply is discussed in greater detail in another section of this paper and in two companion papers.[1,2]

The terminals are designed for operation with E-type in-band signaling for transmission of dial pulses and supervisory signals over the N3 channels. The E signaling equipment is included in the same shop-wired package with the terminal equipment. No provision is made in an N3 terminal for built-in out-of-band signaling such as was used in ON2 and N1 systems. All of the currently available E-type signaling units may be used with N3 carrier channels.

Even though N3 is a superior transmission system with more features and better performance, the price per installed carrier channel will generally be less than that for ON2 carrier channels except for small installations. A substantial price advantage is realized from the shop-wired "packaged terminal" arrangement, from the reduced power requirements of transistorized circuitry, and from reduced space and central office cabling requirements. A more detailed description of the various equipment features follows.

III. CARRIER TERMINAL EQUIPMENT

Except for resistance networks for combining signals and certain equipment which is common to a "terminal package", the equipment for one carrier terminal consists of seventy-one plug-in units. Printed wiring board extensions on the units interconnect to associated equipment when inserted into in-line connectors on the terminal mountings.

Shop-wired strapping and local cable then provide the connections between unit connectors and to the channel patching jacks, E signaling equipment, trunk release and make-busy equipment and fuses in the same package bay. Three important advantages accrue from the plug-in unit arrangement: (*i*) inoperative or malfunctioning units may be easily replaced by spares for prompt service restoral; (*ii*) relatively inexpensive bay equipment may be engineered and installed for future growth and the cost of the more expensive units deferred until later; (*iii*) units needing repair or readjustment may be shipped to repair centers where the work can be done by well-equipped and highly-skilled specialists.

The plug-in units for an N3 terminal are contained in six die-cast aluminum shelves with one shelf cover casting above the top row of units. Each shelf is provided with cast guides for 12 module positions. The shelves are located one above the other and so designed that associated plug-in units are held in place laterally by cast guides on the top of one shelf and the bottom of the shelf above. The shelf cover casting includes the upper guides for the units on the top shelf. The arrangement of units is shown pictorially in Fig. 1 and the shelf and cover castings are shown in Fig. 2.

The plug-in compandor and channel modem units for the 12 channels in channel group 1 are located on the lower two shelves of a terminal. Units for channel group 2 are similarly located on the two upper shelves. The two middle shelves provide space for a plug-in 48-volt to 21-volt dc-dc converter power supply unit, 12 double-channel regulators, two channel group modems, two alarm and restoral units, two frequency correcting units, a group receiver unit, a group transmitter unit, a combining and switching unit, and a line terminating unit. Resistance combining networks for each of the two 12-channel groups are arranged on printed wiring boards attached to a mounting plate. This plate is mounted on the rear flange of the bay between the shelves. Designation card holders are located on the bay upright at the left end of each shelf and are equipped with standard cards having spaces for essential operating company information.

IV. PACKAGED BAY EQUIPMENT — GENERAL

The equipment arrangements for the N3 carrier terminal have been designed to locate the carrier terminal with other closely associated equipment in a shop-wired package. In addition to the carrier terminal, the shop-wired package includes signaling equipment, voice-fre-

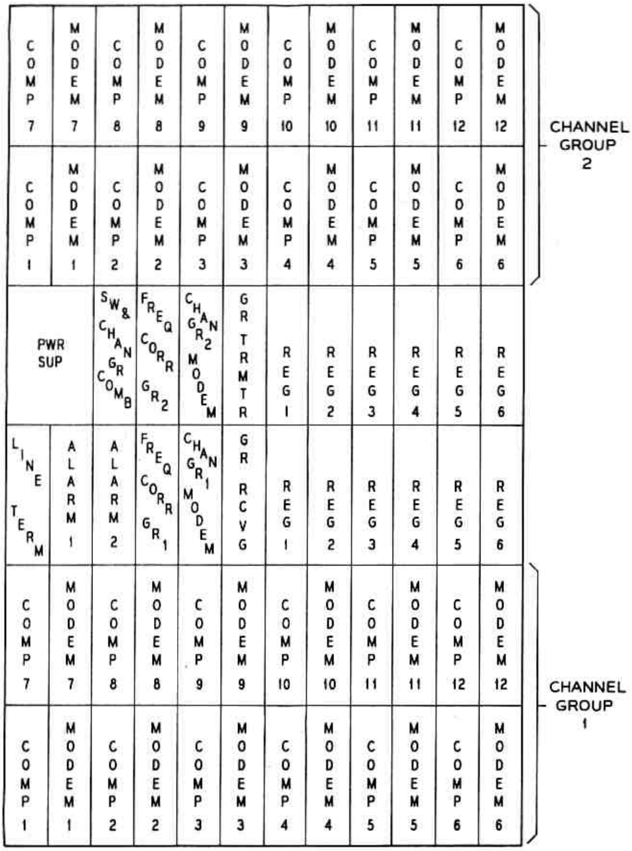

Fig. 1 — N3 carrier terminal — location of plug-in units.

quency patching jacks, and jacks and switches for access to and control of voice-frequency transmission and noise measuring equipment; trunk release and make-busy equipment; battery filtering and distribution equipment; common bay alarm equipment; and miscellaneous common equipment.

Heretofore, the associated major equipment items would have been located in separate bays and interconnected through the distributing frame by cabling run and connected in the central office by the installer. The reduction in central office cabling and connections achieved by this new design concept is clearly demonstrated in Fig. 3. Not only is a substantial amount of installation and engineering work eliminated, but the interconnections made during manufacture are less ex-

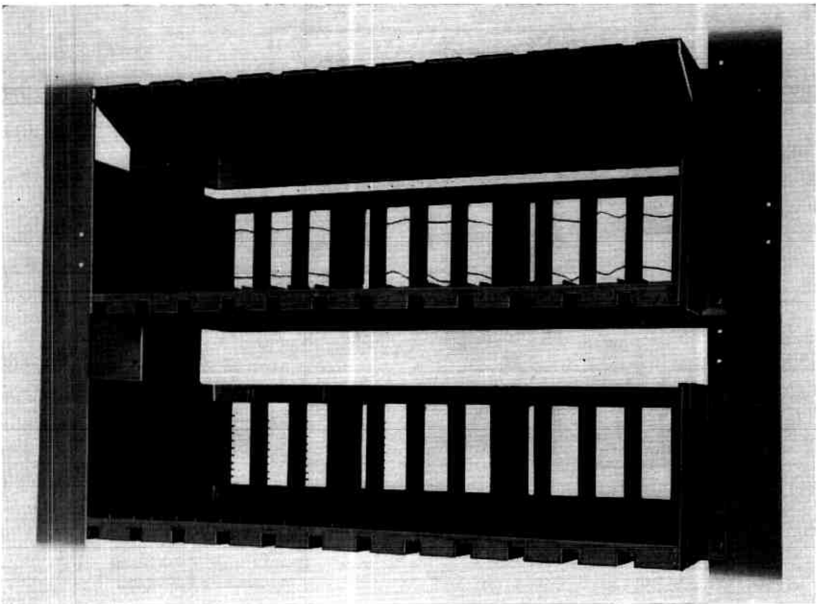

Fig. 2 — Die-cast shelves.

pensive and are completely and efficiently tested by shop testing consoles. The exposure to central office noise and crosstalk pickup is reduced, and the problems of cable rack congestion and segregation of voice and signaling cabling from carrier line cabling are substantially lessened. Existing office arrangements were designed to meet strict resistance limits on some of the E-type signaling leads prior to the development and addition of trunk release and make-busy equipment. Including this new equipment in the signaling path would often require office recabling to stay within resistance limits if the various associated equipments were on separately located bays.

In servicing and maintaining the terminal equipment, there is some advantage in having associated equipment conveniently located together. Permanent association of this equipment also permits related identification with permanent designations for carrier channels, signaling unit positions, and channel patching and monitoring jacks.

A variety of "packaged terminal" frames was designed to accommodate the needs of different central office terminal room arrangements and ceiling heights. The voice-frequency patching and monitoring jacks and associated voice-frequency transmission and noise measuring equipment may be omitted from the package, in the 11 foot-6 inch

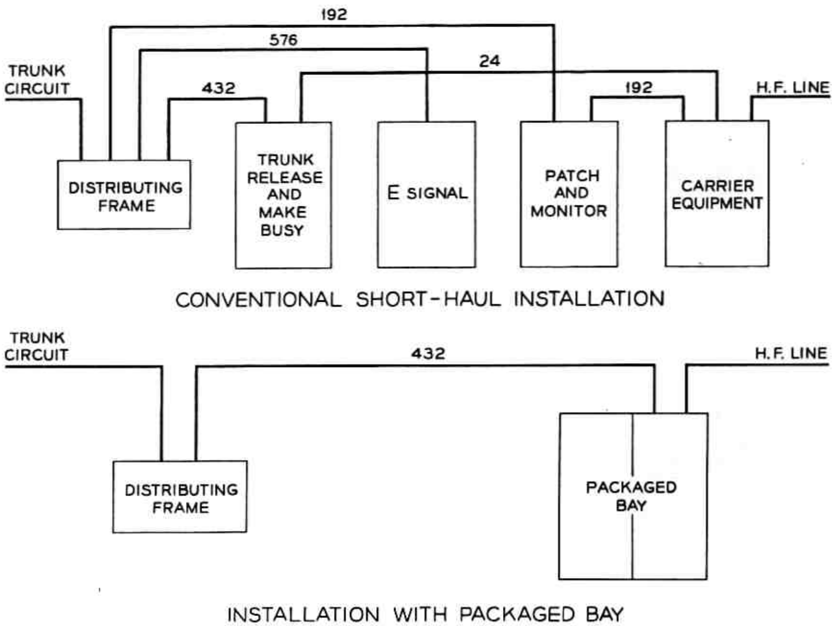

Fig. 3 — Comparison of installer cabling — number of conductors.

bays, where it is desired to locate this equipment in a centralized patching bay instead. Packaged terminal frames are designed in 11 foot-6 inch, 9 foot, and 7 foot heights.

V. PACKAGED 48-CHANNEL BAY EQUIPMENT

The 48-channel package is mounted on an 11 foot-6 inch double-bay cable-duct type framework, which is 53 inches wide and 12 inches deep. This package, shown in Fig. 4, contains the shelves for two carrier terminals, with a monitor and talk panel between them, located in the right side of the framework. The secondary carrier supply panel, from which all carrier frequencies for the two terminals are obtained, is located above the upper terminal. A 2000-cycle tone supply for testing EIF signaling units or a 2400-cycle supply for testing E2B or E3B signaling units is mounted above the secondary carrier supply when required and specified.[2] The power alarm and miscellaneous panel, located above the tone supply, includes power supply fuses for the terminal, signaling, and trunk release and make-busy equipment; alarm relays; means for mounting a restoral oscillator; and a terminal

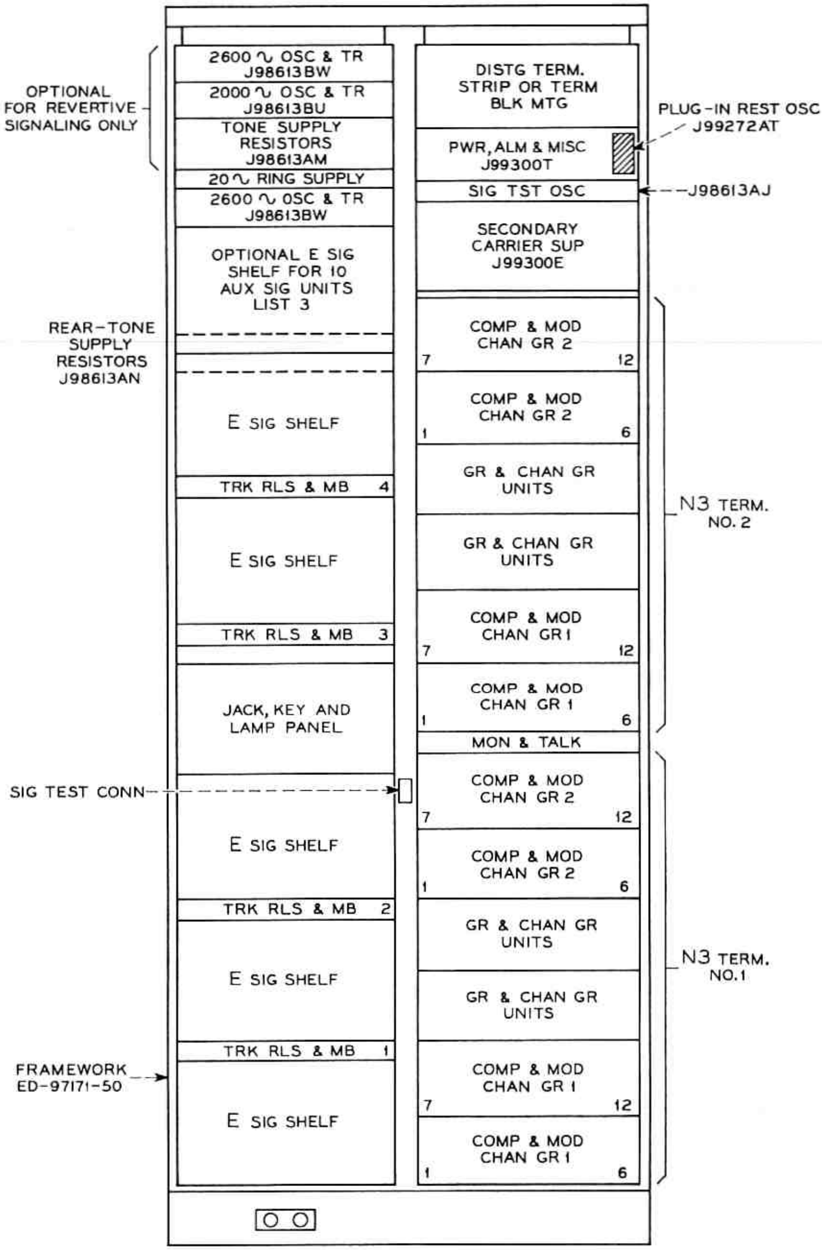

Fig. 4 — J99300A N3 carrier 48-channel packaged terminal frame (11 feet, 6 inches high) with patching jacks.

strip for the carrier line connections and miscellaneous connections to central office equipment. A distributing terminal strip assembly or a terminal block for terminating carrier line pairs may be located in the remaining space at the top of the right side, when required.

In the left side of the double bay, five (a sixth is optional) signaling unit shelves are located for mounting the E-type signaling units. These aluminum die-cast shelves, each of which will mount ten units, are the same parts used in standard E-type signaling bays. Four trunk release and make-busy panels, one for each of the four 12-channel groups, are located between associated signaling shelves. The signaling unit connectors, although mounted on the signaling unit shelves, are shop wired to terminal strips on the trunk release and make busy panel. A jack, key, and lamp panel, which contains the voice-frequency patching and monitoring jacks, is mounted at a convenient height between a signaling shelf and a trunk release and make-busy panel. A 2600-cycle oscillator and transfer panel is located above the signaling unit shelves. Optional equipment, which may be located at the top of the left side, includes oscillator and transfer panels for two frequencies required for revertive signaling. This optional equipment for revertive signaling also includes a tone supply resistor panel. The standard 2600-cycle signaling tone supply is located on the rear flanges of the bay behind the top support of one shelf of signaling units and the bottom shelf of the row of signaling units above it. The bay also includes a signaling tone test connector and an optional 20-cycle ringing supply panel.

All interconnections between equipment in the "packaged frame" are made through shop formed bay local cables. Shielded wire is used for reduction of noise and crosstalk in carrier frequency leads and the more sensitive low level leads are placed in a local cable arm located in the center vertical cable duct. Most of the bay local cable connections are solderless wrapped to achieve manufacturing economy.

An additional 48-channel package is available in which the jack, key, and lamp panel is omitted and a terminal strip assembly is provided through which the voice channels are connected to a centralized patching bay located elsewhere. The few remaining keys and lamps associated with the alarm and control functions and the order wire telephone set jacks are mounted in a jack mounting located between the carrier equipment for the two terminals.

VI. PACKAGED 24-CHANNEL BAY EQUIPMENT

Packaged terminal frames for one N3 terminal (24 channels) are available on 9-foot and 7-foot double-bay duct-type frames. None of the operational features of the 48-channel bay have been omitted. Each frame, as shown in Fig. 5, includes the carrier terminal equipment and units; VF patching and monitoring jacks together with miscellaneous jacks, lamps, and keys; an alarm, power and miscellaneous panel; a secondary carrier supply panel; E-type signaling unit shelves; signaling tone supply equipment; trunk release and make busy panels for each 12-channel subgroup; optional test tone supplies; a 20-cycle ringing supply; and a test connector for signaling test tones and battery supply for portable test sets.

The first 7 or 9-foot packaged N3 frame in an office (or the first of an added pair of frames) must include a secondary carrier distribution

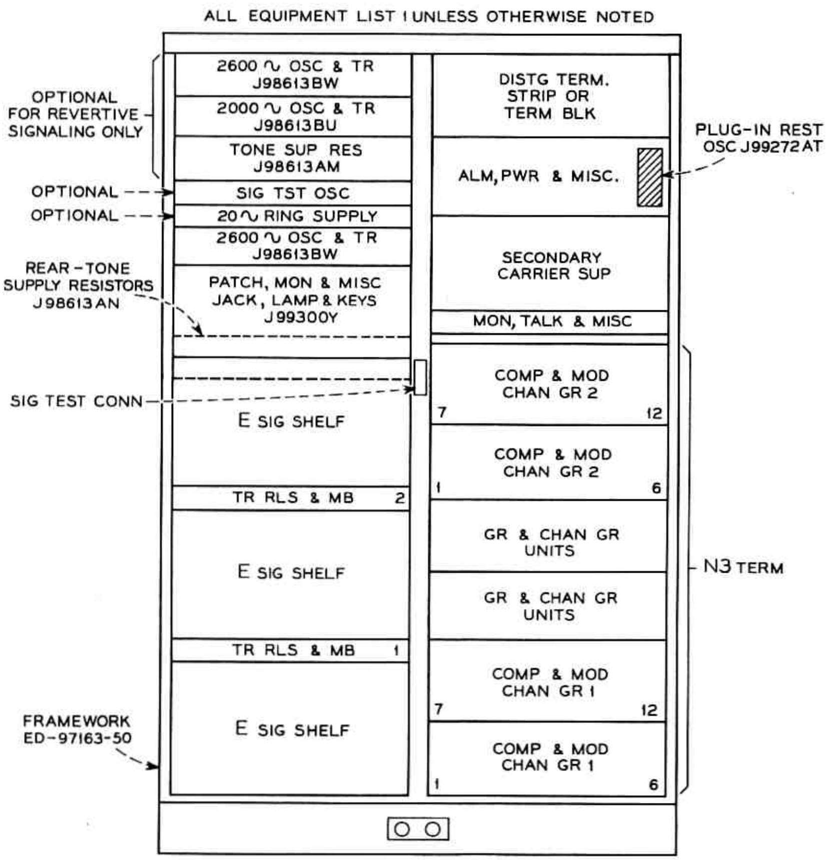

Fig. 5 — J99300D N3 carrier 24-channel packaged terminal frame (7 feet high) with patching jacks.

panel and an alarm, power and miscellaneous panel. Since these two panels are capable of feeding carrier frequencies and power to two terminals, considerable savings per channel can be realized by using this equipment to feed a second N3 packaged frame. For this reason, the second packaged 24-channel frames use terminal strip panels, in place of the carrier supply and alarm and power panels, for terminating interbay cables between the first and second packages.

VII. SIGNALING AND TRUNK CONDITIONING

It will usually be possible to locate all of the E-type signaling units associated with the carrier channels in the same shop wired package with the carrier terminal. Where auxiliary EIL-A and EIS-A units are used, they are matched with companion EIL and EIS units, respectively; each channel so equipped having two signaling units. Where these auxiliary units are used extensively, some additional unit mounting facilities may be required outside the shop wired package.

As noted earlier, the basic 48-channel packaged frames provide for mounting 50 E-type signaling units. An optional shelf may be added within the packaged frame for ten additional signaling units where the use of EIL-A and EIS-A units requires the additional space. In those cases where a total of more than 60 E-type signaling units will be associated with the 48 carrier channels, a signaling shelf is provided at a location outside the packaged frame and installer wired to the terminal strips on the trunk release and make busy panels in the terminal frame as required.

The 24-channel packaged frames provide mounting space for a maximum of 40 E-type signaling units in the 9-foot frame and 30 units in the 7-foot frame. Additional units can be accommodated by use of a supplementary signaling shelf provided at a location outside of the packaged frame.

The signaling shelf positions and associated trunk release and make-busy panels are wired to accommodate all currently existing codes of E-type signaling units. The optional connections for the various signaling unit and trunk circuit combinations are accomplished by wire wrapped connections at the terminal strips on the trunk release and make busy panels. Installer wired cabling to the distributing frame, for the 24 or 48 channels, is also connected to the terminal strips on these panels and includes connections for all optional choices of signaling unit codes. Arrangements are included to bypass the signaling unit connections for direct connection of the carrier channel to the distributing frame, if desired.

Either the 4-wire terminating sets incorporated in some E-type signaling units or separate 4-wire terminating sets located in other bays may be used with N3 carrier channels.

VIII. TRUNK RELEASE AND MAKE-BUSY PANELS

The trunk release and make-busy panel operates under control of the alarm and restoral unit to process the associated trunks during carrier failure and upon subsequent restoral. One panel is provided for each of the two 12-channel groups in an N3 terminal. When a plug-in alarm and restoral unit is provided for each of the two 12-channel groups of an N3 terminal, each trunk release and make-busy circuit operates independently of the other. This arrangement provides maximum circuit protection. Alternatively, one alarm and restoral unit can be used in each N3 terminal to control the two trunk release and make-busy units associated with the terminal. While this latter arrangement does not protect against failure of the channel group modem or frequency correcting units in one of the two 12-channel groups, it will be satisfactory for many applications. For most trunk and line circuits, the functions performed by the trunk release and make-busy circuits are as follows:

(*i*) Conditions the associated trunk or line circuit of an electromechanical switching system to disconnect the busy message trunk circuit, stop subscriber charges, and prevent subsequent trunk seizures during the alarm interval.

(*ii*) Provides alarm indications to a No. 1 electronic switching system which will then make idle trunks busy, disconnect and "make-busy" trunks in the pulsing condition, prevent subsequent trunk seizures, stop charges on calls in progress, release calling and called subscriber lines.

(*iii*) Registers each carrier failure experienced by a channel group on a unidirectional basis; that is, the failure is registered only at one end of a system to facilitate maintenance.

(*iv*) Automatically restores alarms and trunks to normal, at both terminals, under control of the alarm and restoral circuit, when transmission is restored to normal.

The specific circuit functions performed by the trunk release and make-busy circuit for each message channel differ widely in accordance with the E-type signaling unit and trunk or line circuit assigned. To provide the requisite flexibility for channel reassignment without installer effort, separate blocks of terminals on a terminal strip are assigned to each of the 12 channels in a channel group. This arrangement permits reassignment of a channel, which is accomplished with optional straps installed by maintenance personnel, without interrupting service on other channels. A plastic template which can be installed on the block of terminals associated with a channel is provided for each of the channels to indicate the necessary optional strapping required for a particular trunk circuit and signaling unit combination. The template also serves as a visual aid to check the strapping. A wide variety of templates is available covering most of the combinations anticipated for use with N3 channels. Provision of a terminal strip, subdivided into blocks of terminals for each channel, eliminates two distributing frame appearances for each channel. The resultant reduction in office cabling simplifies the problem of meeting critical lead resistance limitations, reduces noise and crosstalk exposures, and achieves substantial economies.

The trunk release and make-busy panel consists of a fabricated steel framework which mounts on the wiring side of the bay and supports the following:

(*i*) A 2-inch mounting panel which is extended forward to the front face of the bay and is equipped with the control relays and a message register.

(*ii*) A terminal strip and associated fanning strip attached to hinged supports at the rear of the framework and arranged to rotate on a horizontal axis into a position which permits access to both sides of the terminal strip and to the wiring side of the relays and message register.

(*iii*) A small terminal strip having the optional strapping which permits the use of either one or two associated alarm and restoral units for each 24-channel terminal.

To facilitate wiring and assembly of shop-wired bays, the panel local cable is wired to E-type signaling connectors for each of channels 3 to 12 of the associated channel group. These connectors are later fastened to an adjacent E signaling shelf when the panel is mounted in the bay. The panel terminal strip provides for connections, through the bay local cable, to signaling connectors for channels 1 and 2 at a more remote location in the bay. Space is reserved on the terminal strip for installer connections to three spare E-type signaling unit connectors and their auxiliaries which are occasionally required and may be mounted external to the packaged frame.

All interconnections between the terminal strip and the relays are included in a shop-wired panel local cable. The bay local cable and installer cabling connect to the normally exposed side of the terminal strip and optional strapping is done on this side. This panel occupies 8 inches of vertical space on the rear flange of the bay uprights, utilizing the space behind an E-type signaling shelf.

IX. MISCELLANEOUS EQUIPMENT PANELS

One alarm, power, and miscellaneous panel is used in each packaged 11 foot-6 inch terminal frame and on alternate 7 or 9-foot terminal frames. This unit includes distribution fuses for the carrier terminals, signaling units, trunk release and make-busy panels, alarm, and miscellaneous equipment in the frame. It also includes the main −48 volt fuse and fuses for feeding +130 volt and −130 volt power over the transmission pairs to remote repeaters. A −48 volt filter is included since the signaling units require filtered battery. The panel provides a mounting and connector for the plug-in restoral oscillator which provides 2600-cycle tone for transmission tests of the carrier line made after a system failure and prior to restoration of service. The relays associated with bay alarm circuits and from which alarm indications are passed to the central office alarm system are located on this panel. The panel also includes a terminal strip for miscellaneous interconnections to equipment in the central office. The restoral oscillator is a plug-in unit and is not furnished as a part of this panel. This panel occupies three and one-half 2-inch mounting plate spaces and is mounted and wired in the shop.

A patching, monitoring, and miscellaneous jack, key, and lamp panel is provided in one version of the packaged 48-channel equipment and in the 24-channel packages. This panel is mounted and wired into the frame and provides the following:

(*i*) The 4-wire VF channel patching and monitoring jacks.

(*ii*) Transmission and noise measuring jacks, keys, lamps, and switches.

(*iii*) The 1000-cycle test supply jacks.

(*iv*) Interbay patching and testing trunk jacks.

(*v*) Alarm lamps for transmission failure indication and an alarm override key and lamp for each 12-channel group.

(*vi*) Alarm lamps and alarm release keys for power supply unit voltage alarms.

(*vii*) Jacks and keys associated with the 4-wire monitor and talk circuit.

(*viii*) Telephone set jacks for association with the N carrier order wire.

This equipment is arranged on a jack panel approximately 23 inches wide and either 10 or 13 inches high for the 24- and 48-channel packages, respectively. Permanent designations are stamped on the panel and designation strips are provided where operating company circuit designations will be required.

X. DOUBLE-BAY DUCT-TYPE FRAMEWORK

The newly developed double-bay duct-type framework, shown in Fig. 6, is used for all terminal packages. Its use, instead of single bay

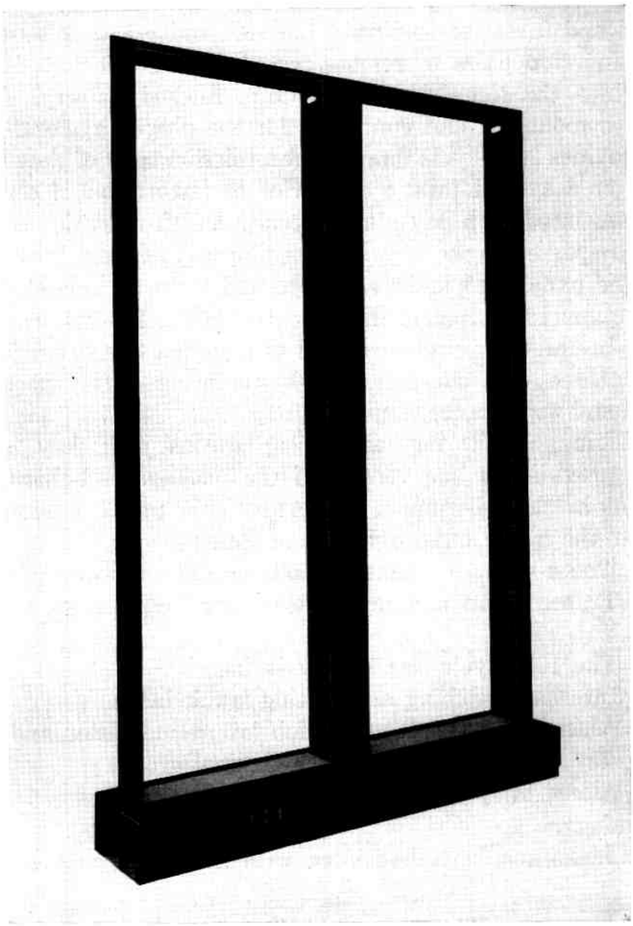

Fig. 6 — Bay framework (front view).

framework, permits more efficient utilization of space and greater economy in the shop assembly and wiring of a package. A number of equipment items which must be in a package of any size can serve the double bay at no greater cost than for a single one.

This bay has five-inch deep uprights with wide flanges in front and narrow flanges in the rear. The uprights with their flanges form cable ducts at the sides and middle of the frame. Both front and rear flanges are drilled and tapped for mounting bay equipment. The narrow rear flanges afford access to the duct for shop wiring or installer cabling. With this arrangement, low-level leads can be placed in the middle duct where they are automatically shielded from noisy or high-level wiring in adjacent bays. High-level wiring and wiring not too susceptible to noise is located in the outside ducts of the same frame. The wide front flanges increase the duct space and also provide sufficient space for designation card holders on the uprights.

The bases of the frames include guard rails and commercial power outlets on both front and rear for ac supply to testing and maintenance equipment. The bases and upper sections of the frameworks are designed for ready attachment of tools to facilitate handling in the shop and in the telephone office during installation. The frame is also arranged for attachment of installers dolly trucks which permit wheeling the bay into place in an upright position.

XI. TERMINAL PLUG-IN UNIT DESIGN — GENERAL

Since it was anticipated that the demand for N3 carrier plug-in units would be relatively large, it was decided that the equipment should not be designed around existing hardware and apparatus if significant penalties were imposed on the design as a result. It appeared that development costs involved with new hardware and apparatus could be justified if the new designs would reduce manufacturing or maintenance costs per unit even slightly. Working closely with the Western Electric Manufacturing Organization, considerable effort was expended to make the designs as adaptable as possible to automated machinery.

The basic plug-in unit consists of a single-sided, glass epoxy, printed wiring board mounted in an aluminum die-cast frame. Fig. 7 shows a typical board, frame, and shield in their relative final positions.

The shield is mounted to the frame directly adjacent to the wiring side of the board so as to provide a ground plane between adjacent

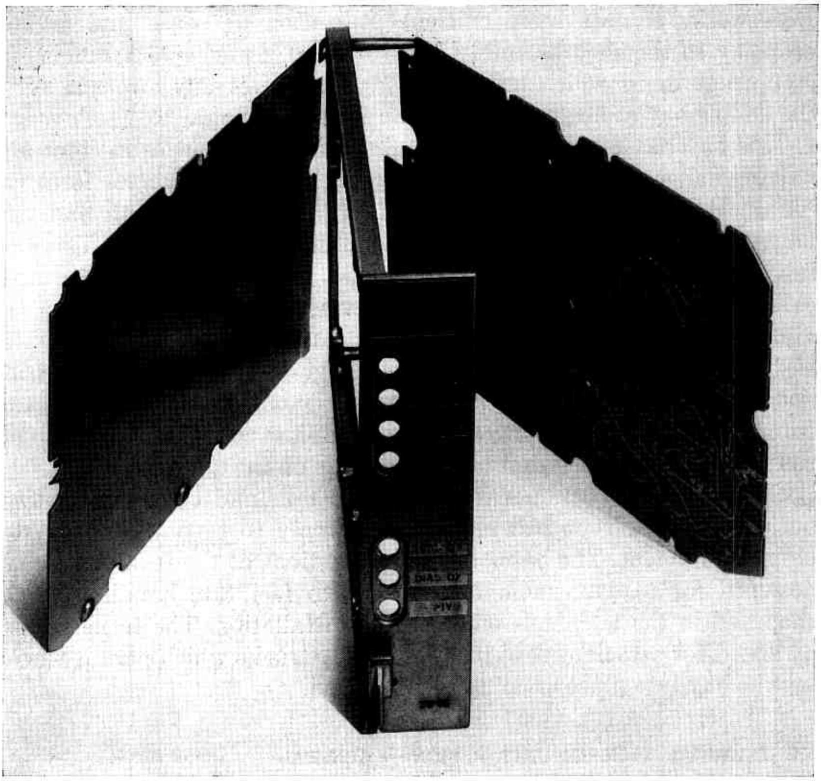

Fig. 7 — Basic frame, printed wiring board, and shield.

units for signal isolation. The shield is made from either aluminum or permalloy, depending on the amount of isolation required.

Both the printed board and shield are mounted to the frame with hex washer head thread-forming screws. These screws have a triangular shaped cross section designed to mount into cored holes in the frame. As the screw is installed, a thread is rolled in the frame without the formation of chips common to most self-tapping screws. The screws maintain a high residual torque after installation so that locking devices are not required. The washer head provides a fine bearing surface for mounting the board and shield without damaging them.

The frame die casting is of a very simple design, requiring very little post-casting machining because of the fact that no tapped holes are required. The basic frame has two surfaces on which the board and shield mount. The top and bottom frame rails provide the track designed to slide into the bay shelf casting. The faceplate and rear posts of the casting provide protection to the components on the board when the assembly rests on the component side. The faceplate on the card frame extends the full width and height of a plug-in position and so provides a neat appearance when a shelf is filled with plug-in units. The modular space required by this unit is 8 inches high, 12 inches deep and 1¾ inches wide; thus, twelve units may be plugged into a single 23-inch shelf casting. The faceplate of the casting is provided with cored access holes to test points and potentiometers for in-service testing and maintenance. Each unit has a different faceplate due to different test points and adjustable components which appear there. These castings are all made by one basic die having interchangeable faceplate slides.

The only machining operation required on the basic casting is the sawing of a groove in the faceplate and lower frame rail for mounting a latch. Fig. 8 shows many of the pieces of hardware developed specifically for N3. Item 6 shows the latch assembly complete with spring and pivot pin. Although the latch resembles other latches used in plug-in equipment, the design of this latch is such that it is impossible to remove a plug-in unit without operating the mechanism. The angle of the hook is such that forces tending to dislodge a unit apply pure tensile forces to the latch, free of any twisting moment about the latch pivot point. The latch engages a hole in the shelf casting. When the latch is operated, the latch hook is released from the hole in the shelf and, in addition, the unit is pried forward to disengage the plug on the unit from the connector mounted in the bay.

Surface wiring is not only costly to provide but is a common source of manufacturing problems. To eliminate as much surface wiring as possible, any test point, potentiometer or switch which must be accessible at the faceplate is mounted directly to the printed wiring board at the faceplate edge of the board. Items 7, 1, and 9 of Fig. 8 show such parts which were designed for the N3 system.

The test point, shown as Item 7 on Fig. 8, was developed for use in the N3 carrier system. Two electrical connections from the test point are made with the printed wiring board by tabs which are clinched and soldered to a wiring path. To keep from stressing these electrical connections during probe insertion, two plastic tips protrude through the board and are headed over on the wiring side. The barrel of the test point is made from beryllium-copper, ¼ hard so as to provide a reliable contact. The tabs of the connector are made from relatively soft brass, welded to the barrel, so as to allow clinching without frac-

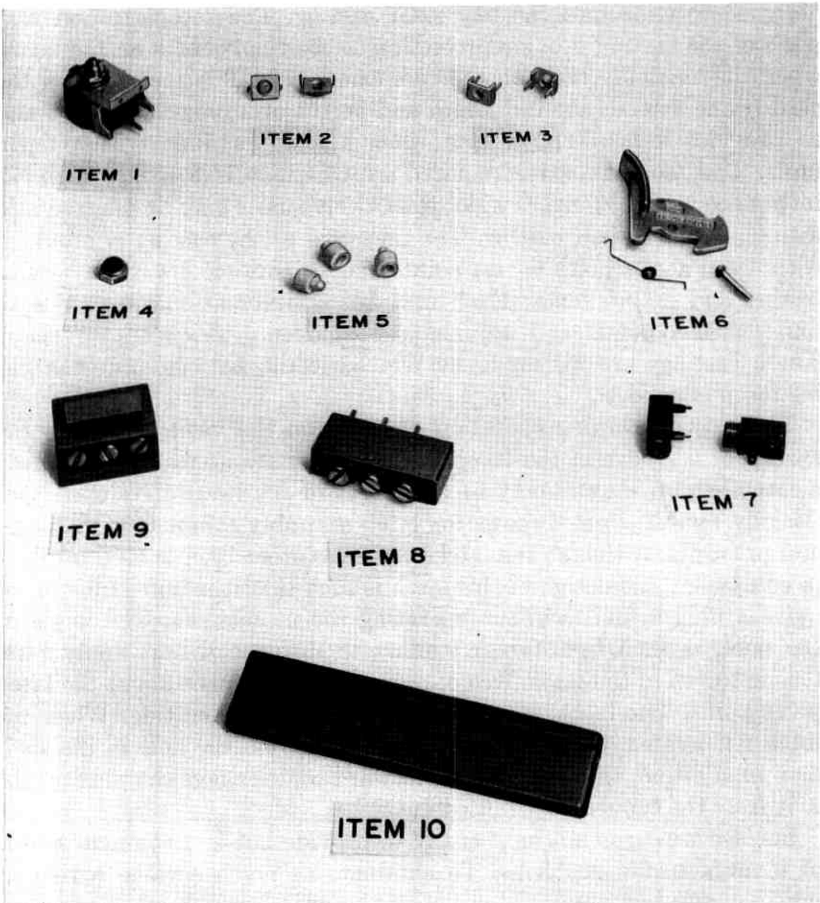

Fig. 8 — Plug-in unit hardware.

ture. The plastic surrounding the barrel of the test point protrudes through holes in the faceplate so that the test probes will not short to the casing during testing.

The potentiometer, shown as Item 1 of Fig. 8, was designed for N3 use. Access to this adjustment is obtained by reaching through a hole in the unit faceplate with an ordinary screwdriver.

The switch, Item 9 of Fig. 8, was designed to provide a component which, like the test point, mounts to the printed wiring board and protrudes through the faceplate. The switch presses into a cut-out on the front edge of the printed wiring board. Electrical connections are made via the wires which protrude through the plastic block and bend so as to pass through the printed board where they are soldered to land areas. Item 8 of Fig. 8 is the same basic type of switch as Item 9 except it is made to mount internal to a unit directly on the printed wiring board without front access.

The filters and equalizer networks used on the N3 plug-in units are all housed in cans and are moisture resistance sealed where necessary. Two separate methods of electrically connecting these networks are employed — soldered type and mechanically attached type. The soldered type has electrical terminals and studs located on the mounting surface of the can. The terminals and studs protrude through holes in the printed board. When installed, the electrical terminals are soldered to land areas and the filters are mechanically held with elastic stop nuts mounted on the studs. The nuts generally bear against electrically grounded printed land areas so as to provide an electrical ground connection to the studs and, hence, to the cans.

Mechanically attached networks are those which are intended to be mounted or easily changed for maintenance in the field. For these units, it is desirable that no solder connections be required and that they be mounted without removing the metal shield from the card frame. Items 2 and 3 of Fig. 8 show piece parts designed to serve the function of a nut, tied electrically to the circuit, yet accessible from the component side of the printed board. Item 2 mounts on the wiring side of the board while item 3 mounts on the component side of the board. Both nuts are clinched to the board and soldered to land areas on the wiring side. The method in which they are used will be shown later.

The shield has such a large span and is located so close to the wiring side of the board that some provision to keep the shield away from the solder connections is necessary other than mechanical fastening to the frame at its edge. Items 4 and 5 of Fig. 8 are two parts designed for this purpose. Item 4 is an elastic stop nut which is used for mounting soldered type filters. The nut has a nylon insert which extends high enough to provide the necessary clearance between the board and shield, the nylon surface being higher than any solder mass. Where filters are not used, a nylon part (item 5) is snapped into a hole in the board from the wiring side to provide an insulating spacer.

With the exception of the power supply unit, all N3 plug-in units employ a printed wiring type of connector. The male portion of the connector consists of gold-plated printed wiring contacts on a portion of the printed wiring board which extends beyond the frame casting in the rear. Item 1 of Fig. 9 shows both the connector portion of the

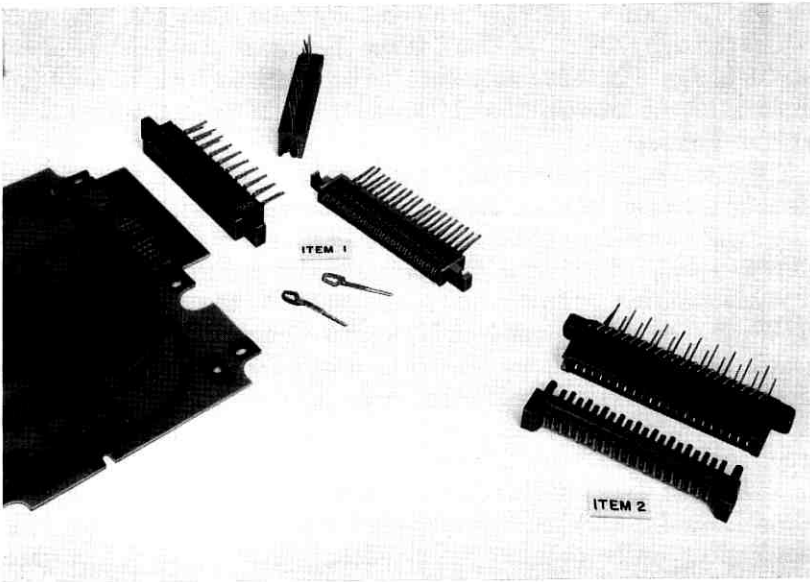

Fig. 9 — Plug-in unit connectors.

printed wiring board and the connector which mounts in the bay. As can be seen from the various cut-away views of the female connector, each tab has two pretensioned surfaces which wipe against the conductor path for a reliable connection. At the point where the surfaces touch the path, precious metal buttons are welded to the spring. This precious metal, wiping against the gold-plated land area, provides very low contact resistance. These connectors are mounted to the bay with a specially designed shoulder screw which provides the necessary float for alignment.

The power supply, due to its large current flow, requires a heavier contact than that provided by the printed wiring connector. For this purpose, the two part connector shown as item 2, Fig. 9 was used. The item 1 connector was designed specifically to meet the needs of the N3 carrier system.

Item 10 of Fig. 8 shows a flexible plastic protective cover which is placed on the connector portion of the printed wiring board before leaving the manufacturing area. This throw-away cover protects the gold-plated tabs from damage during handling and shipment. Since the connector tab portion of the printed wiring board is not protected by the frame, glass epoxy material was chosen for the basic board because it withstands impact without damage much better than phenolic materials. The use of glass epoxy material was also found to be necessary for resistance to shock and vibration, since the apparatus mounted on these large boards is sometimes heavy and produces large stresses on the boards when vibrated or shocked.

The module height of the basic plug-in unit was chosen to allow the compandor, modem, and double-channel regulator units (large volume production units) to be of simple design for low manufacturing cost. In these designs, all components are placed on the main board. Some of the group units had too great a component density to allow all of the components to be placed directly on the main board. In these cases, a slave board, Fig. 10, is used. Two piece-part clips were designed to hold the edge of the slave board and provide stability. The clip is soldered to the slave board and also, after clinching, to a printed area on the main board so that the clip can serve as a conductor path connecting these boards electrically as well as mechanically. The components are mounted to the slave board in the manner shown in Fig. 10. Note that the pigtails of the components themselves are bent in a manner such that they serve as the electrical connection between the slave board and main board.

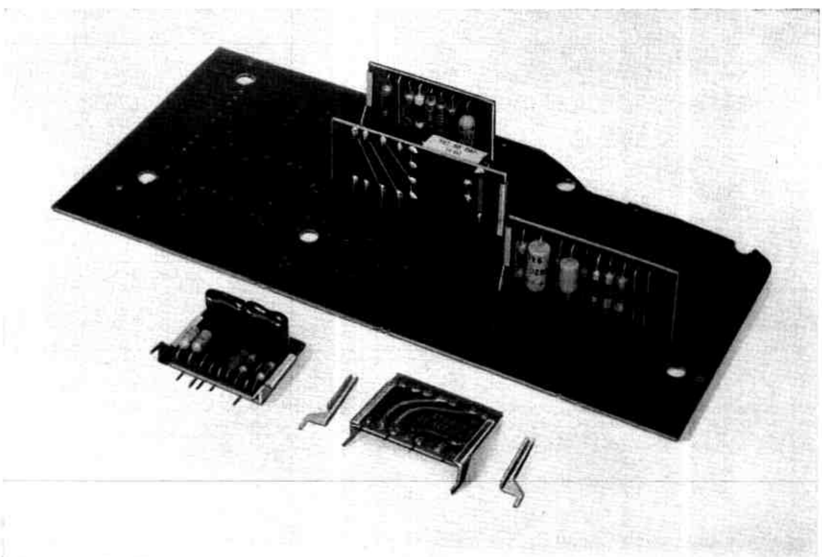

Fig. 10 — Slave board technique for obtaining high component density.

XII. SPECIFIC PLUG-IN UNIT DESIGNS

Several of the plug-in units will be discussed. In some cases the discussion will use specific units merely as examples to discuss general design principles whereas in other cases units will be discussed because of their unique design features.

Fig. 11 shows the component side of an early model of an N3 compandor unit. Some of the special hardware items developed for N3 are easily seen. The test points and potentiometer which mount to the board, accessible from the faceplate, are located at the faceplate end of the board. Terminal ※1 of the connector is externally wired to a frame ground. To connect frame ground to the unit frame and shield, the board mounting screw in the lower right-hand corner of each plug-in unit mounts through a specially developed terminal which is clinched and soldered to connector tab ※1. The mounting screw, bearing against this terminal, transfers the ground from the connector to the frame and from there to the shield via the shield mounting screws.

The compandor circuit, containing a compressor and expandor circuit, is laid out on the board so that there is adequate electrical isola-

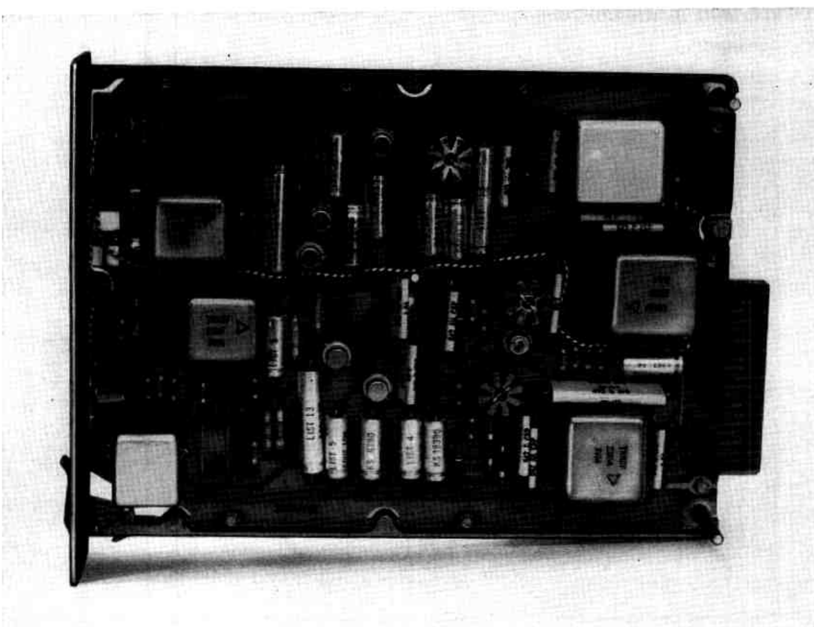

Fig. 11 — Compandor unit.

tion between these two circuits. Not only are the two functions physically separated but paths carrying low-level signals in the compressor circuit are spaced as far as possible from paths carrying high-level signals in the expandor circuit and vice versa. This principle was adhered to on all units having two-way circuits. Even the ground path common to both portions of the circuit was kept separate on the board and made common only at the connector end of the board. Some units even have completely separate ground paths so as to prevent any possibility of local currents causing crosstalk interference.

Note that all of the components of the same type are mounted on the same bending centers in order to reduce the number of insertion machine positions required in the manufacturing area. Most tubular components lie in the same plane not only to make insertion by present type machines relatively easy but also to make the job of insertion compatible to fully automated machinery if and when it is used.

Two twisted pairs of surface wire were required in the compandor in order to meet stringent requirements on crosstalk. In many cases the wiring pattern became so involved that paths were required to cross on the printed wiring side of the board. This crossing is done with pieces of bare wires which are treated exactly as components from a manufacturing point of view. All such bare wires are of the same length so that they may all be inserted by the same machine. Although small details such as this may seem quite trivial, they are really very significant in view of the fact that these units will be manufactured in great volume.

The channel modem unit, Fig. 12, demonstrates the two basic types of filters and networks used on N3 plug-in units. The two filters are the bandpass filters required to select a specific 4-kHz signal from the spectrum of 12 channels, one filter for each direction of transmission. The filters are completely attached to the unit from the component side of the board since these filters may be installed or replaced in the field. Since each of the 12 modem units requires different filters, a single modem unit, less filters, is provided to the field along with the required filters. This principle is also used for the pick-off filters of the double-channel regulator unit. The mechanical mounting of these filters is accomplished with the specially designed nut discussed previously. Screws having captive lock washers are used for mounting the filters to assure a good electrical connection between the filter and the path to which the nut is soldered. Since these filters contain very sensitive crystals, and inductors operating at very high impedances, they are moisture sealed. Any attempt to provide a method of attaching

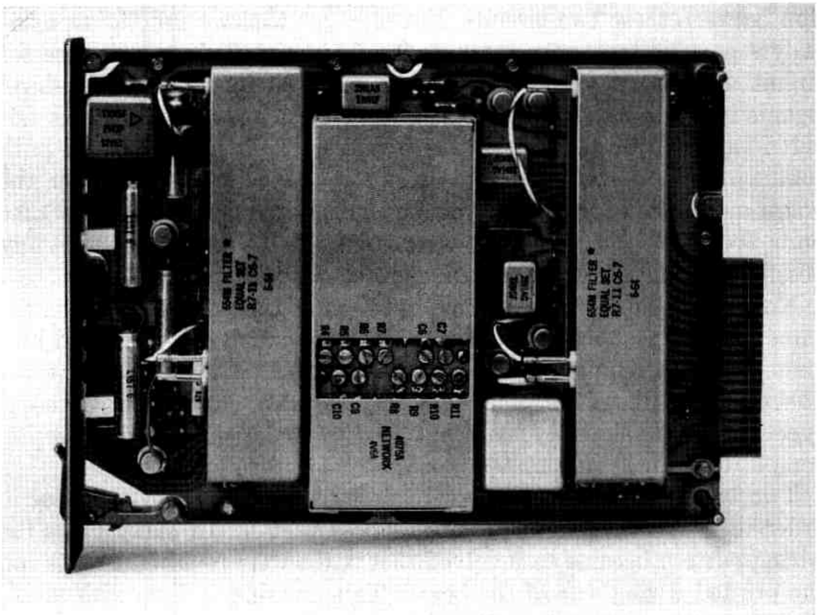

Fig. 12 — Modem unit.

leads to the electrical terminals which puts excessive stress on these terminals has to be avoided in order to maintain the seal integrity. A study of the tolerances involved in designing hardware for this purpose indicated that it would be best to use flexible wiring for the connections. But, since soldering was not desired, a special gold-plated connector was coded which would be soldered to a wire from the board and would be pressed on to the filter terminals. Gold-plated connectors and filter terminals assure low contact resistance.

The network located in the center of the board is an equalizer network which works in connection with the bandpass filters to achieve the desired bandpass characteristics. Since there are twelve filters of different frequencies used on channel modem units, the equalizer characteristics may have to be different for each of the twelve channels. All of the required components are placed in this single can which is soldered into the board during manufacture. The installer matches the equalizer to the particular set of filters used by turning down the prescribed set of shorting screws. The particular screws which require turning down are designated on the filters.

The high group receiving unit, Fig. 13, demonstrates the complete

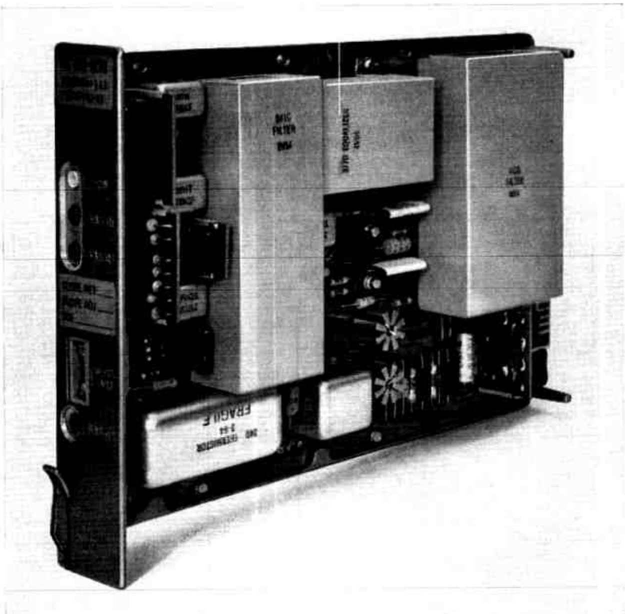

Fig. 13 — High group receiving unit.

absence of wires on the component side of the board. The slope adjust switch and test points are all mounted on the printed wiring board and protrude through the faceplate so as to be accessible to the craftsman without removing the unit from service. Four slave boards are used to provide high component density. The slave board is treated just as any other multileaded component.

The smaller of the three cans on this board is an equalizer providing partial slope equalization for the N line. Since line frequency characteristics vary, this equalizer must be selected by the transmission engineer to match the line. For this reason, this can is designed to be mounted in the field rather than in the manufacturing shop. Unlike the bandpass filters used on the modem unit, this equalizer is not a moisture sealed unit and it was possible, by using the two special nuts described previously, to mechanically and electrically mount the can to the board with screws.

The alarm and restoral unit, power supply unit, combining and switching unit, and line terminating unit are shown in Figs. 14, 15, 16, and 17, respectively. These units are shown not because they represent typical N3 plug-in units but rather because they have features peculiar to themselves.

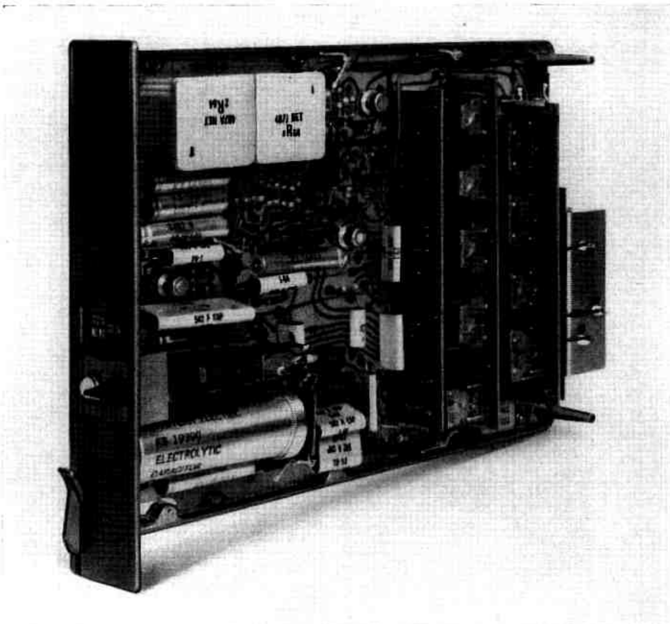

Fig. 14 — Alarm and restoral unit.

Fig. 15 — Power supply unit.

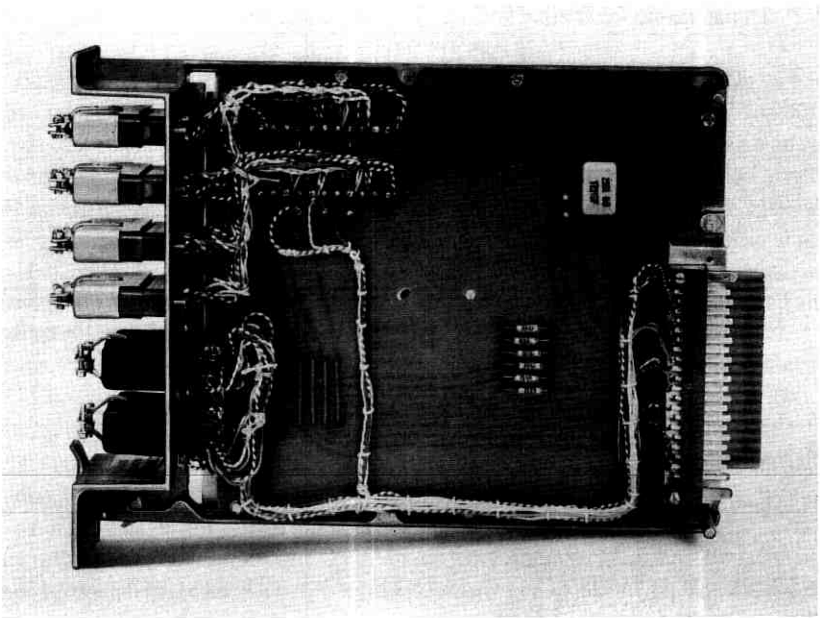

Fig. 16 — Combining and switching unit.

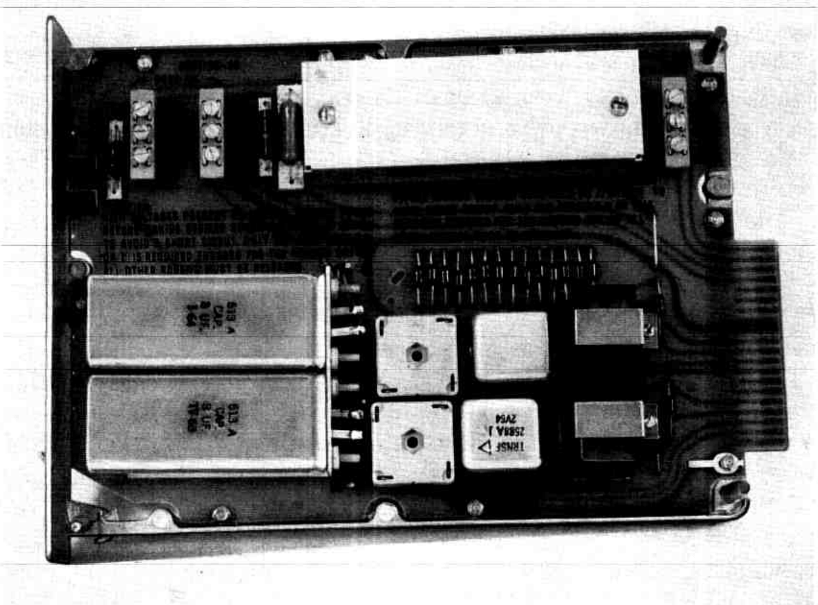

Fig. 17 — Line terminating unit.

The alarm and restoral unit, Fig. 14, uses 9 miniature relays. The amount of cross-connecting wiring required between relays was so great that many wires were required in addition to the printed paths. Slave boards, containing only bare wires instead of components, provided a method of running paths across the main board on the component side without using surface wires. A large amount of power is dissipated in one resistor on this unit. The resistor is mounted in a U-shaped metal shield at the rear of the unit so as to radiate the heat out the back of the bay where it will have negligible effect on the temperature rise of the unit. The large electrolytic capacitor, key, and lamp, although usually mounted directly to the casting, are mounted in this unit on special brackets which mount directly to the printed wiring board.

The power supply unit, Fig. 15, is shown because of its unusual construction as compared to the other N3 plug-in units. Since this unit required two modular spaces and since its components are rather large and bulky, it is made of a fabricated housing with most of the components mounted directly to the housing. The regulator circuit is partially contained on a printed wiring board mounted within the housing. This unit requires many loose wires due to the type of construction. Since these wires are contained within the housing, they are protected from damage related to handling.

The combining and switching unit, Fig. 16, uses two connectors for connecting to the bay wiring, one of the printed wiring type and one of the type used on the power supply unit. The resistors and hybrid coil on the printed wiring board provide the combining and splitting functions in the transmitting and receiving paths of the group units. Six female connectors, equipped with male shorting plugs, are located on the faceplate of this unit. Two connectors are provided for each of the group transmitting, group receiving, and power supply units. These connectors are used in conjunction with the switching set to permit in-service replacement of degraded group and/or power units.

The line terminating unit, Fig. 17, is shown to demonstrate how line treatment (build out, simplexed power feed, and surge protection) is accomplished. The two plug-in plastic cases in the lower right-hand corner are span pads to provide the required loss to the transmitting and receiving cable pairs. The screws, three sets of three each, provide the necessary means of sending power down the line for sealing non-soldered cable splices or powering repeaters from different local battery voltage supplies for different repeatered line requirements. The large heat shield contains a power dropping resistor which, depending on the amount of simplexed power required, may dissipate between 0 and 11 watts. This heat is directed away from the unit and towards the shelf castings so as to keep the temperature rise of the unit to a minimum.

XIII. SWITCHING SET

The switching set (Fig. 18) is used in connection with the combining and switching unit to provide in-service switching of either a group transmitting unit, group receiving unit, or a power supply unit. Since these units can become degraded without complete failure and since they handle 24 voice channels, it is desirable to be able to replace them with new units without interrupting service. Upon removing one of the two paralleled power connector shorting plugs on the combining and switching unit, a power cord from the switching set is plugged into the

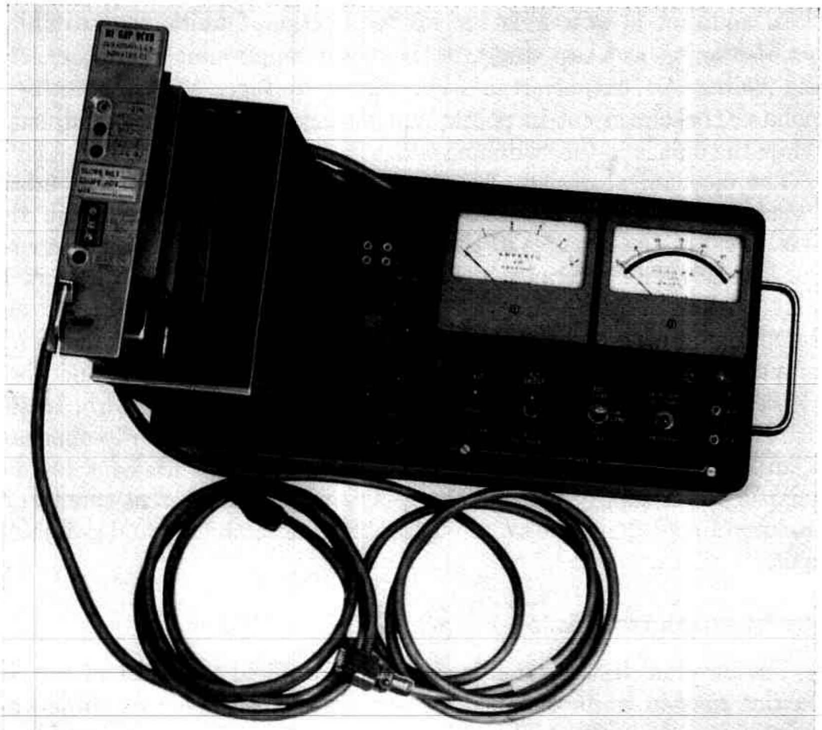

Fig. 18 — Group and power unit switching set.

vacated connector to supply power to the switching set. An alternate unit of the type to be replaced is plugged into connectors in the switching set (in Fig. 18, a high group receiving unit is shown being plugged in). To switch a group unit, a transmission cord from the switching set is plugged into the appropriate group unit connector on the combining and switching unit from which a shorting plug has been removed. After equalizing the output signals of the alternate and in-service units by use of gain controls on the switching set, the second shorting plug on the combining and switching unit may now be removed so that the bay service can be switched to the alternate unit. The regular unit may now be replaced by a new unit which is switched into service by reversing the procedure used to take it out of service. To switch a power supply unit, the procedure is similar to that used for switching a group unit except in this case the switching is accomplished by increasing the voltage of the alternate unit until it assumes the full load of the bay equipment. By switching in this manner, negligible "hits" will be experienced.

In addition to switching the above mentioned units, the switching set also serves as a test stand for the power supply unit to allow checking voltage of output, regulation, ripple voltage, and alarm cut-in points. The alarm cut-in points and voltage output may be adjusted while the unit is in the switching set.

The mechanical design of the switching set is a very simple fabricated box construction with all of its circuitry contained within the box. Most of the features of the switching set are obvious by referring to Fig. 18. The alternate unit wiring is not accessible when it is plugged into the set. This is a desirable feature in order to prevent tampering with the circuit when it is in service. The switches which actually transfer the load from the main to the alternate unit and vice-versa are protected to prevent their being accidentally thrown. In the same way that the power supply unit required a "heavier" connector than the other plug-in units, a "heavier" connector is used for switching the power supply unit which explains why two different connectors are used for plugging the switching set into the combining and switching unit.

XIV. TERMINAL TEST STAND

The terminal test stand, Fig. 19, is used to field test most of the N3 carrier plug-in units on an out-of-service basis. The two cords and printed board plugs are placed in the bay to extend the bay connectors down to the connectors in the test stand. Two cords are required since

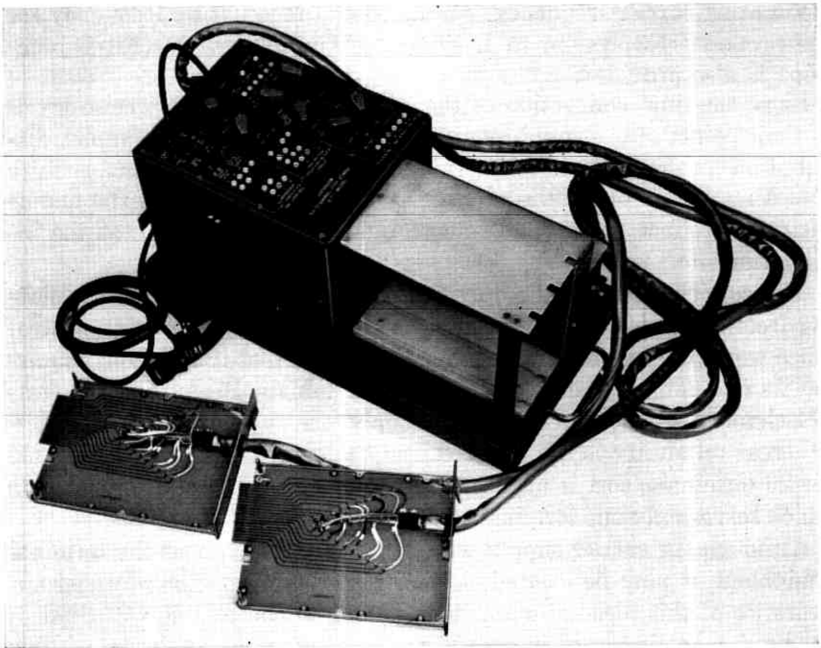

Fig. 19 — Plug-in unit terminal test stand.

some tests require two different units to be tested at the same time. The actual tests which are performed on the test stand are too numerous to be discussed in this paper. Each conductor of the test stand cords has a pin-type test jack appearance available on the stand for trouble shooting.

The terminal test stand is designed to be functional but not extravagant. Since the demand for the stand is estimated to be small, it is made completely of fabricated parts. As in the case of the switching set, the wiring for the test stand is completely enclosed within the body.

XV. CARRIER SUPPLY — GENERAL

The N3 carrier telephone frequency supply is common equipment that furnishes the modulating and demodulating frequencies for a maximum of 26 N3 carrier terminals. This requires the generation of 12 frequencies for voice-to-carrier frequency modulation of 24 voice channels, two channel group modulating frequencies and a group modulating carrier frequency. An additional modulating frequency for the proposed N3 system to L system group conversion (N3-L junction) is also provided.

Since the final connection of the generated carriers at precise levels to their points of use could require wiring of 1482 pairs, a complex distribution problem is encountered. Economic considerations require that a maximum of the distributing connections be made in the manufactured product with as few connections as possible made at the installation site.

The carrier supply is the only source of modulating and demodulating frequencies for an N3 terminal. It must be used in a small central office where only a single terminal is used as well as in locations requiring its maximum capacity of 26 carrier terminals. Because as many as 624 channels might depend on this supply, the system must be reliable. Whereas reliability is the prime consideration of the supply when used for 26 terminals, cost is of major consequence to the small installation which serves only one terminal.

Although the carrier supply will, in most cases, be near the terminal equipment, it may be located elsewhere in the office. The physical appearance of this piece of equipment should be compatible with that of the equipment it serves in order to aid the plant forces in associating the carrier supply and the terminal. This continuity of appearance is established by using the same die-cast unit frame and mounting shelf as used in the terminal equipment.

As outlined above, distribution, reliability, cost and appearance are the major design considerations in common equipment. The following description represents a solution to these problems for the N3 carrier supply, the mechanical arrangement of which is shown in Fig. 20. This figure indicates the relationship of the prime carrier generating and distributing equipment to the secondary distribution and carrier terminal bays.

XVI. FREQUENCY GENERATION

All of the frequencies distributed from the carrier supply are generated from a 4-kHz source.* Two different optional plug-in units provide the 4-kHz drive to the saturating inductor which creates the harmonics of 4 kHz. One plug-in unit contains a stable, temperature-controlled crystal oscillator (coded the 61A oscillator) for independent generation of the 4-kHz signal. The second unit is driven from a 4-kHz source

* A detailed circuit description is included in this issue by R. L. Haner and I. E. Wood, "Circuit Design of the N3 Carrier Terminal."

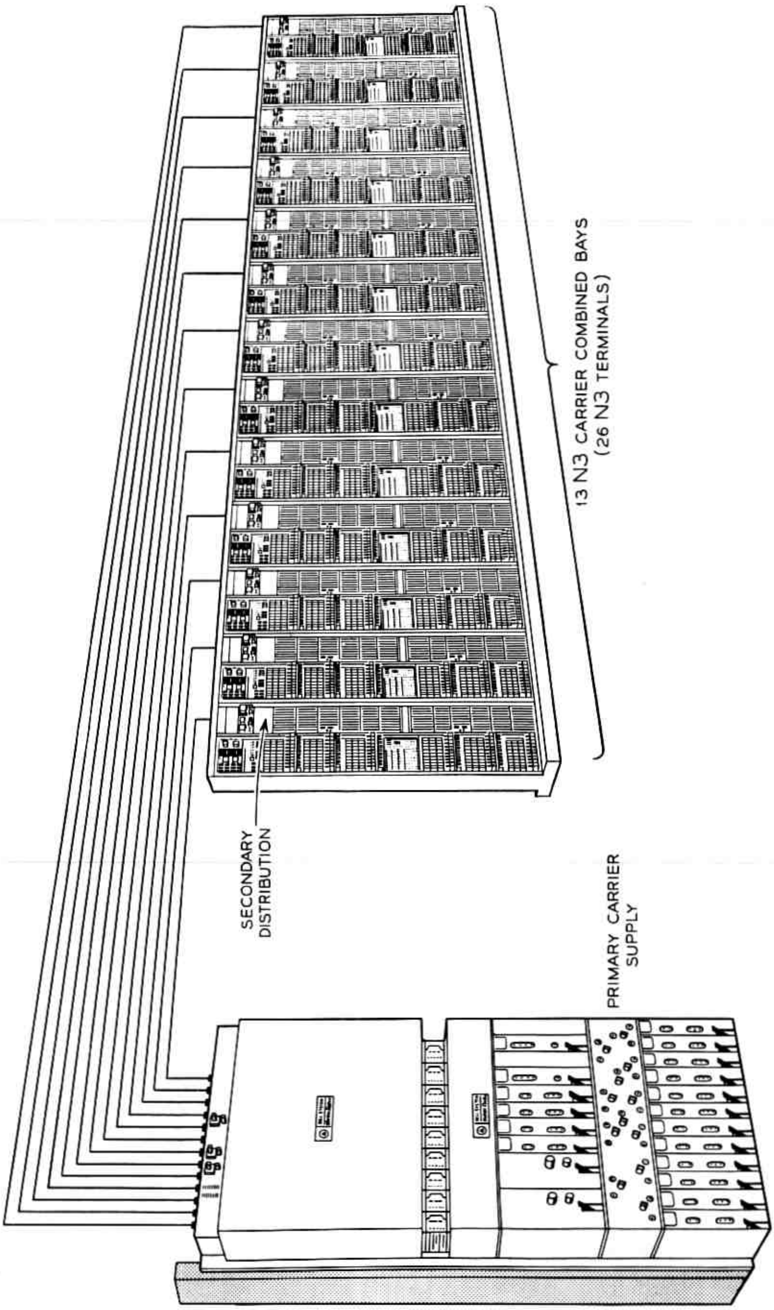

Fig. 20 — N3 carrier supply in relation to the carrier terminals it serves.

already available in the office. The harmonic generator creates the even and odd harmonics of 4 kHz in the required carrier range. Fifteen of these harmonics are selected by crystal filters. A sixteenth frequency for group carrier modulation (304 kHz) is generated in the doubler-amplifier from the 152-kHz tone selected by one of the 15 crystal filters. The filtered harmonic outputs are applied to limiting amplifiers which provide a constant output signal voltage determined by the regulated power supply voltage.

The 4-kHz generator plug-in units are constructed on a basic N3 modular plug-in unit frame with an additional modular face plate attached to provide the necessary height from the printed wiring board surface. Such an arrangement is necessary to accommodate the polyurethane plastic foamed oven of the 61A oscillator on the unit which provides independent generation of 4 kHz. Each type 4-kHz plug-in unit contains sensing circuits which activate alarms and control circuitry in case of excessive variations in oven temperature (when provided) or output voltage of the 4-kHz signal. Fig. 21 shows the two versions of the 4-kHz generator, (a) receives 4 kHz from an external source and (b) contains an internal 61A oscillator.

The harmonic generator and crystal filters necessary to select the required frequencies are positioned together on a single panel. The harmonic generator is driven by the 4-kHz generator and provides even ($(2n)$ 4 kHz) and odd ($(2n + 1)$ 4 kHz) products on two separate balanced output circuits. The odd and even filters are separated by being fastened to opposite sides of the mounting shelf to simplify the wiring of the assembly. Fig. 22 shows the 4-inch high housing, with and without its front maintenance panel, indicating the equipment arrangement of the completely enclosed harmonic generator and filters. All connections to this unit are made through a terminal block located at the back of the unit. External wiring from the unit to the associated equipment is run in shielded pairs to reduce office noise pick-up into these low-level circuits.

The frequencies developed by the harmonic generator and filters are applied to the input terminals of the dual amplifier plug-in units shown in Fig. 23. The dual amplifier units contain two identical voltage-limiting amplifiers. A built-in level sensing circuit monitors the output of both amplifiers and provides an output alarm indication if either amplifier is out of working limits. Two frequencies selected on the basis of circuit reliability and intermodulation considerations are applied to each dual amplifier and separately amplified through them. Seven dual amplifiers provide amplification of 14 frequencies by am-

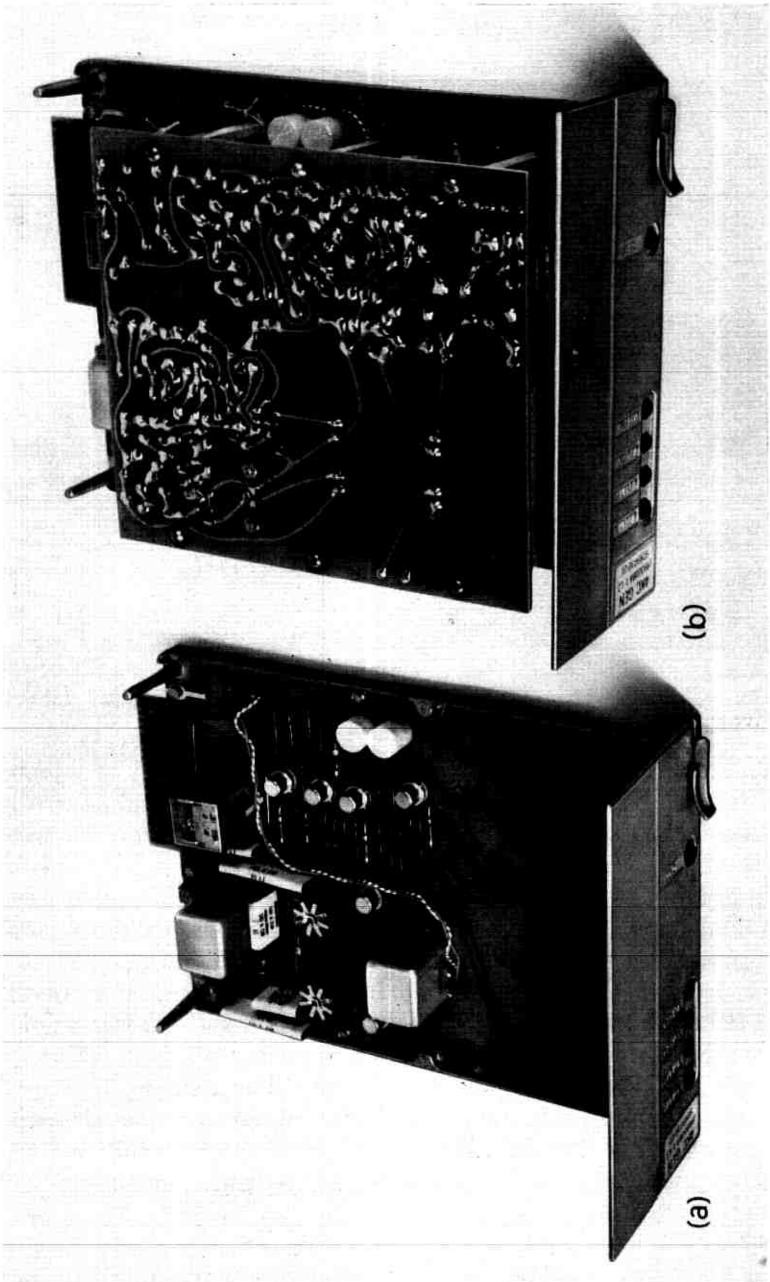

Fig. 21 — 4-kHz generator plug-in unit. ((a) for use with external source; (b) with 61A oscillator mounted in place.)

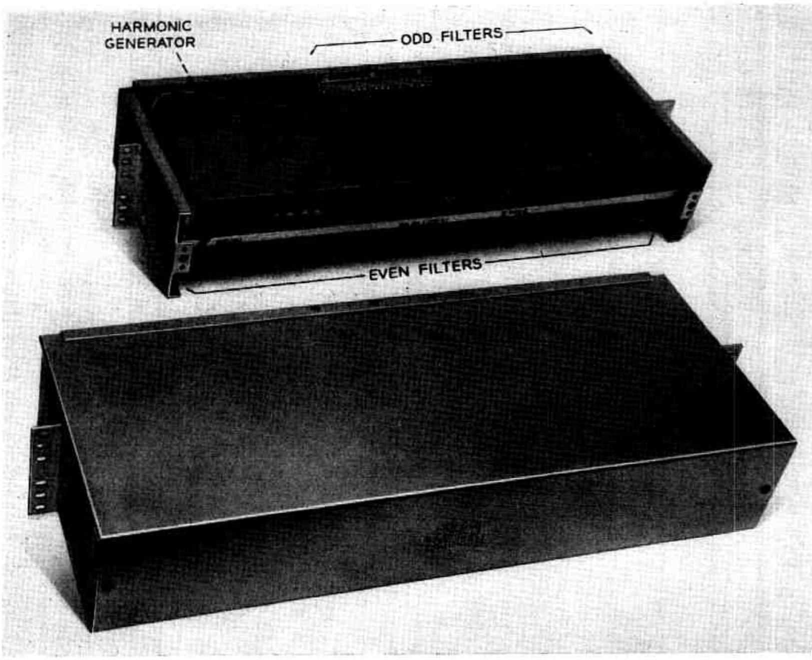

Fig. 22 — Harmonic generator and carrier filter panel shown with and without protective cover.

plitude limiting amplifiers and these outputs are applied as square waves to the primary distribution panel. The fifteenth and sixteenth frequencies (152 kHz and 304 kHz) require different treatment. The 152-kHz signal is applied to a doubler-amplifier, shown in Fig. 23, which is a plug-in unit containing a printed wiring board identical to that used in the dual amplifiers. The doubler-amplifier contains the two amplifiers that the dual amplifier contains; however, a doubling stage is provided in front of one of the amplifiers and the entire circuit is driven by 152 kHz only. The 152-kHz signal is amplified and appears on one output; it is also applied to the doubler stage to generate a 304-kHz signal which is amplified and appears at the second output.

All of the plug-in units are assembled on N3 die cast modular plug-in unit frames. The dual amplifiers are identical to each other and are completely interchangeable from one frequency position to another.

XVII. RELIABILITY

With the carrier supply providing carrier frequencies to many different terminals, the loss of 624 channels (which its total failure would

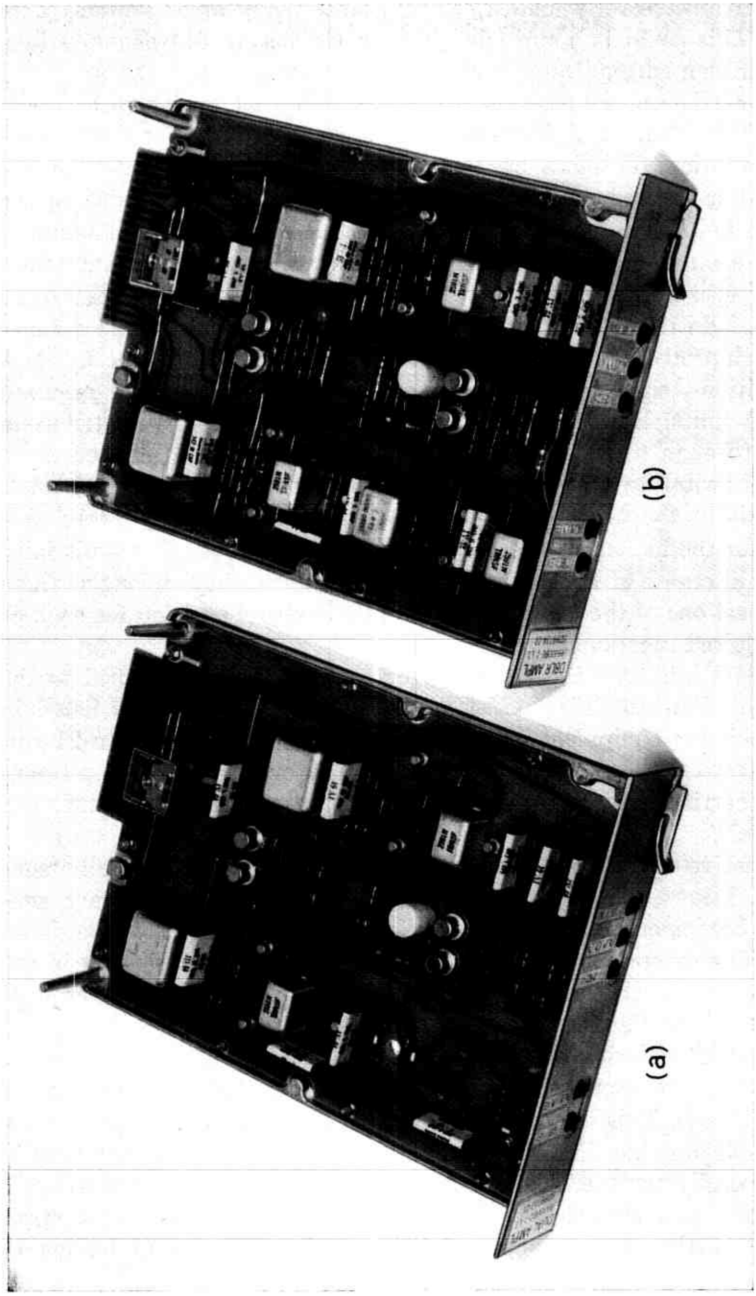

Fig. 23 — (a) dual amplifier; (b) doubler amplifier.

cause if it were fully loaded) could isolate many areas served by an office. This could be more adverse than the loss of 624 channels in a large office might indicate.

The carrier supply has been designed to allow as much choice as possible in the amount of protection to the carrier system for any particular situation. This has been accomplished in two ways. All of the alarm sensing and most of the logic for switching is included in the plug-in units themselves. The sensing features of each circuit required to operate the office alarms in case of failure (even though automatic switching is not required) is a simple arrangement that makes it possible to add the switching logic at very little cost. The second manner in which protection is accomplished is by providing two adjacent plug-in positions, one regular and one alternate, for each of the required plug-in units. Both positions are wired to the primary distribution panel through a control relay. Either position can be selected by a manual switch to have its output connected to the distribution panel; the unit in the other position acts as the automatically switchable spare for the manually selected unit in the event that this circuit fails. Either alternate or regular positions can be left empty as long as there is at least one of the correct plug-in units in either position for each of the required functions.

Table I indicates the pairs of frequencies that are assigned to the different dual amplifier positions. In addition to the units listed in this table the equipment is powered by a plug-in 48 volt dc to 21 volt dc converter whose failure would mean the complete loss of all generated carriers. This is the same power supply used in the N3 packaged terminals.

All active circuit components are mounted on the plug-in units mentioned. Filters, wirespring relays, and inactive components have generally been mounted on fixed panels and hard wired into the signal path. The exception to this rule in the entire carrier supply is in the secondary distribution panel (mounted in the packaged terminal frame) where the regulators for the transmitted carriers are active circuits that are wired in place.

A certain measure of reliability is attained merely by the use of plug-in units. This is not to say that because a unit is plug-mounted it is more reliable. Rather, even if more prone to failure, the plug-in unit is easily replaced and the circuit down time is therefore reduced. Of course, this premise is based on having adequate spares for the protection of the carrier supply. Admitting the necessity of having at least one spare of each type of plug-in component of a supply for re-

TABLE I — FREQUENCY ALLOCATION TO PLUG-IN UNITS

Unit	Frequency	Channels Affected on Single Terminal	Channels Affected on 26 Terminal
4-kc Generator	24 kHz	24	624
Double-amplifier	152 kHz 304 kHz	14* 24	338 624
Dual amplifier	168 kHz 280 kHz	14* 12	338 312
Dual amplifier	172 kHz 232 kHz	2 12	52 312
Dual amplifier	176 kHz 256 kHz	4* —	52 —
Dual amplifier	148 kHz 180 kHz	2 2	52 52
Dual amplifier	156 kHz 184 kHz	2 4*	52 52
Dual amplifier	160 kHz 188 kHz	4* 2	52 52
Dual amplifier	164 kHz 192 kHz	2 4*	52 52

* Since those frequencies marked 4 are transmitted carriers a loss of these carriers would affect 4 channels or 2 per channel group. Half the channels go open and the other half would have excess gain due to the operation of the double channel regulator. Similarly, the 152 and 108 kHz affect 12 channels in one group and 2 in the other.

liability, the spares can be stored in the alternate circuit positions so that they can be powered and their outputs monitored continually for defects and automatically switched to protect the working circuits.

Table I shows that a single spare of each type of plug-in unit, plus a power supply unit, can give automatic protection for most of the critical frequencies. Since the power supply is identical to that used in the terminal equipment, it can be stored in the carrier supply as an automatic switchable spare and be available on a plug-in basis in case of failure of a power supply in the terminal equipment. Likewise, the dual amplifier spare can be stored in the alternate position for 168 kHz and 280 kHz, as an automatically switched spare, and still be available on a plug-in basis for any other dual amplifier position in case of failure. When a unit is used in this manner the originally protected position does not have the switchable protection until a good unit is plugged into the spare position again. Whether a unit is in a regular or alternate plug-in position, it is constantly monitored by its own sensing circuits. By plugging the spares into their automatically switched protection positions their status is always known by the display of indicator lamps.

The condition of plugged-in units is visually presented on a 4-inch switching and alarm panel. The relation of this panel to the plug-in equipment is shown in Fig. 24. In the figure, the regular (or minimum) number of plug-in units are discriminated from the optional alternate plug-in positions which are shaded.

The switches and lamps associated with each set of regular and alternate plug-in positions are located in proximity to the units they control and monitor. This places a lamp either directly below or above the particular unit whose condition it indicates. The switch is placed above or below and between the two units it controls and indicates the manually selected units by means by a line inscribed on the switch knob.

Lighted lamps indicate failure of a plug-in unit. If a unit is not plugged into position, the lamp associated with that plug-in position will not be illuminated. Likewise, if all units are plugged in and are in good working condition, the lamps will not be lighted.

Failure of a protected unit will result in an automatic switch to the protecting unit. The lamp associated with the failed unit will be lighted, and a minor alarm will be indicated. If even a momentary defect occurs in the unit in the position selected by the manual switch, it will cause switching to the protecting unit and will lock the selected unit from returning to active service. This provision is provided to prevent chatter switching of marginal units. The position not selected by the manual switch does not have this lock-out feature and alarm indications occur only as long as a unit is in a failed condition. A minor alarm will be sent, however, and will stay on until it is reset even if the trouble clears. This could create a problem if a marginal unit were located in a protecting position. Service would not be affected, however, and a close watch on the equipment would indicate the defective unit. The alternate position serves as an excellent check of the units capability to function in the circuit since the lamp associated with it will indicate its condition when the unit is plugged in.

If a unit which has no adjacent protecting unit fails, its associated lamp will glow and a major alarm will be indicated. The unit will not try to switch to the other position. This is true whether the other position is blank or contains a defective unit. If both units fail, the unit remaining connected to the circuit will be the one indicated by the position of the manual switch.

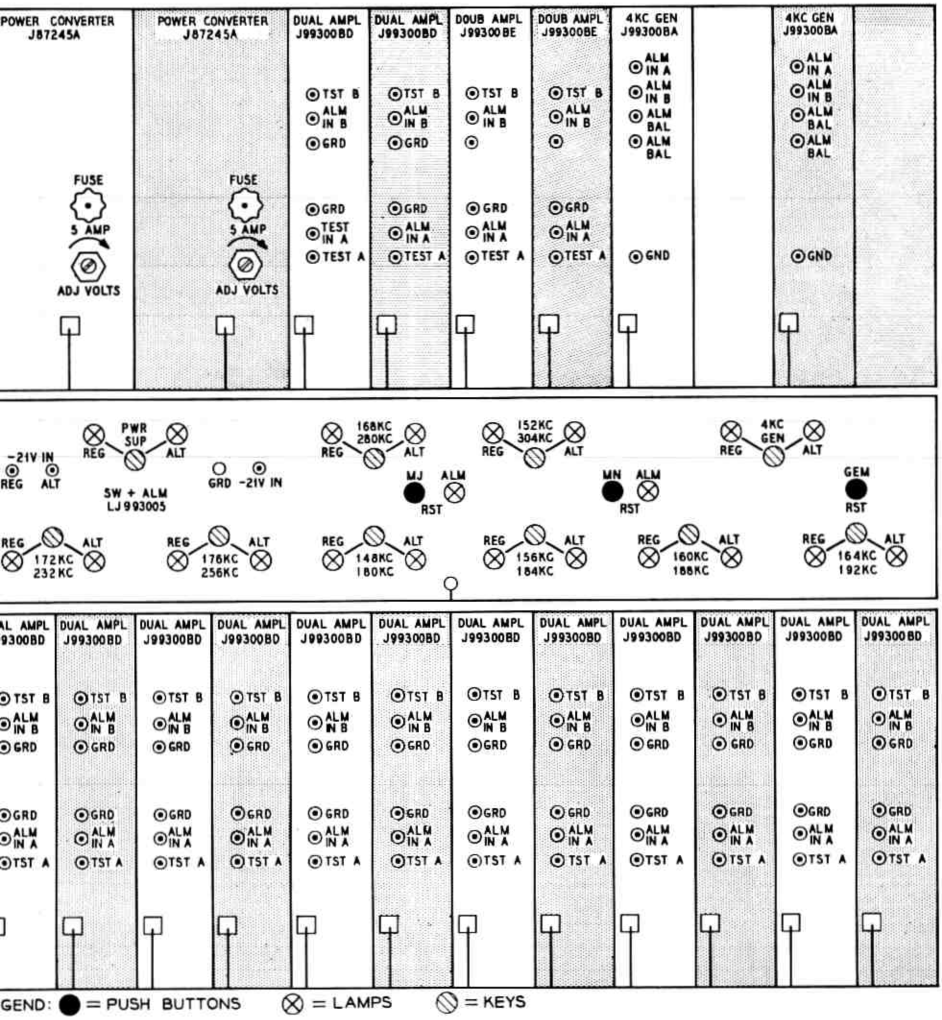

Fig. 24 — Switch and alarm panel in relation to plug-in units (alternate plug-in positions shown shaded).

All of the above functions are equally applicable to units in either the regular position or the alternate position; however, there is a preference for operating with units in the regular position. If -48 volt battery supply to the alarm circuit fails, all of the regular positions will be connected to the distribution circuit regardless of whether they are equipped with working units or not. Automatic switching of all of the units except the power supply is done by transfer contacts on wire-spring type relays. Units not in the circuit are terminated in resistive loads. The power supply is switched by a mercury contact relay and the protecting power supply is always connected to a 50 per cent load to maintain its regulated output voltage within the required range to allow automatic switching.

The major and minor alarm indications are generated by bistable logic circuits which are triggered by an ac-coupled signal from a failed unit. An office alarm indication may be cut off at the switch and alarm panel prior to clearing the trouble condition on the panel itself by operation of the manually operated reset button. After being reset, the major and minor alarm circuits are capable of providing another office alarm for any additional trouble. The logic circuits for the major and minor alarms are packaged plug-in cards. Since they are switching circuits and are not as subject to component aging as analog circuitry and since they are not in "transmission" path, they are available for maintenance only by some disassembly of the switch and alarm panel.

XVIII. DISTRIBUTION

To supply 16 discrete frequencies from the primary supply to the exact location of each carrier terminal where they are needed entails a distribution of 57 pairs to each terminal as shown by the frequency disposition described in Table II.

Two N3 terminals are supplied in an 11 foot-6 inch framework;

TABLE II — FREQUENCIES REQUIRED FOR OPERATION OF ONE N3 CARRIER TERMINAL

Freq. (kHz)	Channel Mod.	Channel Demod.	Trans. Cx	Freq. Correct.	Channel Group	Line Group	L-N3 Group
143	2	2					
152	2		2	1			
156	2	2					
160	2		2				
164	2	2					
168	2		2	1			
172	2	2					
176	2		2				
180	2	2					
184	2		2				
188	2	2					
192	2		2				
232					1		
280					1		
304						1	
256							4

therefore, the frequency distribution in a single bay requires wiring of 114 pairs. In addition, those pairs supplying carrier frequencies which will be transmitted over the carrier line require severe distance limitations between the terminal and the carrier supply or a mop up at the terminal since their output power is required to be within ± 0.05 dB of the normal operating power of -19 dBm.

In order to relieve wiring congestion, deliver the desired frequencies within the amplitude tolerance necessary, and relieve the installer of the large number of connections between the carrier supply and the terminal, a secondary distribution panel is provided in the terminal bay. This distribution position provides compensation for cable wiring losses between the primary and secondary distribution points. It also provides a place within the bay to fan out a single input of each frequency received from the primary distribution to the number of output pairs required for two terminals. Regulation of those frequencies for use as transmitted carriers is also done at this secondary location with the resultant advantage of lower losses in the shorter lengths of distribution leads through bay wiring to the terminals. A secondary carrier distribution panel which provides these features is shown in Fig. 25.

With this distribution arrangement the installer has only to connect

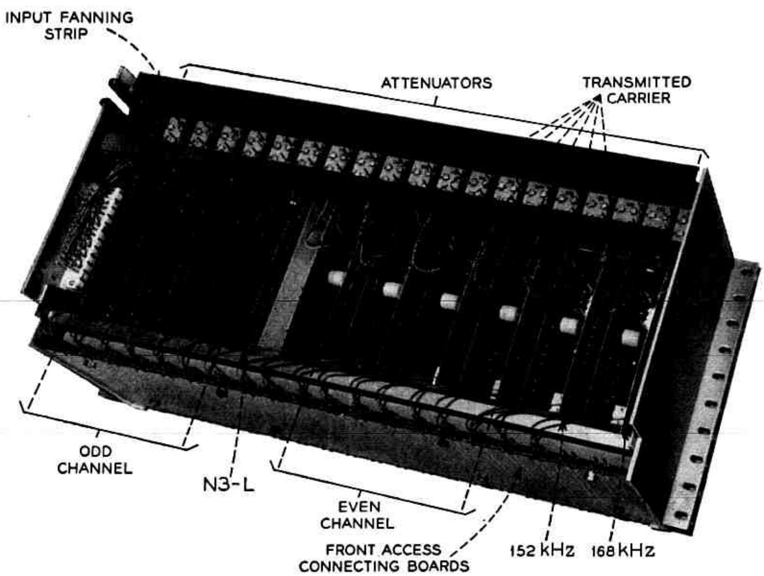

Fig. 25 — Secondary carrier distribution panel (rear view).

a single cable containing 26 pairs (of which 19 are used) between the primary distribution and the secondary distribution panels. All other distribution connections are made on shop-wired bays at the factory.

The secondary distribution equipment consists of a fabricated aluminum chassis which holds four die castings that provide the top and bottom of two shelves. The two die castings consisting of the top and bottom of the upper shelf are identical and provide evenly spaced slide positions for 19 printed circuit attenuator cards. The attenuator cards contain switchable balanced pads that provide 1.5 dB of total loss in 0.5-dB steps for compensation of office wiring losses. Pad selection is made by screw-down shorting adjustments on a plastic block mounted as a component on the attenuator card. The block is mounted to give installer access to the screws without removing the printed card. Connections to these printed cards are accomplished through solderless wrap terminals which are mounted on the edge of the cards and soldered to the printed wiring paths.

The installer clamps the office cable from the primary distribution unit to the side of the secondary distribution panel and dresses the pairs through a plastic fanning strip to the attenuator card locations. He then wraps each pair of conductors to the input terminals of the attenuator provided for each specific frequency. Each of the channel group frequencies (232 kHz, 280 kHz) and the group carrier frequency (304 kHz) are provided on two pairs from the primary distribution so that these frequencies are provided with two attenuators and distributed directly from the attenuators to the terminal circuits.

Although the die castings for the top and the bottom of the lower shelf are identical, the printed wiring boards for which these castings provide slides contain apparatus of different height, and the slide positions are not evenly spaced. The printed boards mounted in these slides provide the final distribution and isolation circuits for two terminals.

There are four different types of boards in the lower shelf to provide the required functions. Seven boards contain capacitors and fixed pads for distribution and isolation of eight outputs of each of seven frequencies. Four outputs of six of the boards are wired to each terminal and provide all of the odd channel ($(2n + 1)$ 4 kHz) modulating and demodulating frequencies. In addition, one of these boards provides four outputs to each terminal for the N3-L junction frequency.

Four boards contain capacitors and fixed resistor pads and provide four outputs of each of four frequencies. Two outputs of each of these boards provide modulating frequencies for the even ($(2n)$ 4 kHz) channels except for 152 kHz and 168 kHz. Two cards with six outputs of 152 kHz or 168 kHz provide the two outputs per N3 terminal for modulating at these frequencies and a third output per terminal for the frequency correcting circuit.

The only active circuitry in the distribution network is contained on six boards with four outputs each for use as transmitted carriers. A regulator and isolating pads are mounted on each of these printed wiring boards.

A total of 19 boards slide into the lower shelf. Shields made of preformed aluminum fit into slides between the printed wiring boards to decrease the crosstalk coupling between adjacent cards.

Connections to all of the boards are made through edge-mounted solderless wrap terminals which are soldered to the printed wiring paths. The terminals are positioned on the boards such that the terminals across the shelf line up in numbered rows and by board position. Ducts provided for fanning the connecting wire hold the boards from coming out of their slides. A duct can be loosened and moved out of the way in case a board has to be removed for repair.

All factory and installer wiring is on the installer aisle side of the equipment. The front, or maintenance aisle side of the equipment, consists of a blank surface 4 inches behind the front guard rail. Two printed wiring boards mounted under the secondary distribution panel provide an appearance of each modulating frequency at the front of the panel. It is anticipated that when arrangements are developed for alternate message and wideband data use of a portion of the carrier line bandwidth, the circuit which accomplishes the required switching may be mounted on the front surface.

Since the secondary distribution panel is mounted in the same bay as the terminal equipment and performs all of the necessary level control and distribution of frequencies to the terminals, the primary distribution has only to distribute the carrier frequencies to a maximum of 13 secondary distribution panels. The connection between a primary distribution panel and a secondary distribution panel is made by a 26-pair, aluminum shielded, polyethylene covered cable. Crosstalk coupling between the cable pairs and pair loss at the higher frequencies limit the separation between a primary distribution and a secondary distribution to 700 sheath feet of this cable. Provision is made for grounding the aluminum shield at either end of the distribution cable by means of a special clip.

The primary distribution panel shown in Fig. 26 has two basic functions. The first and more obvious is to provide for isolated distribution

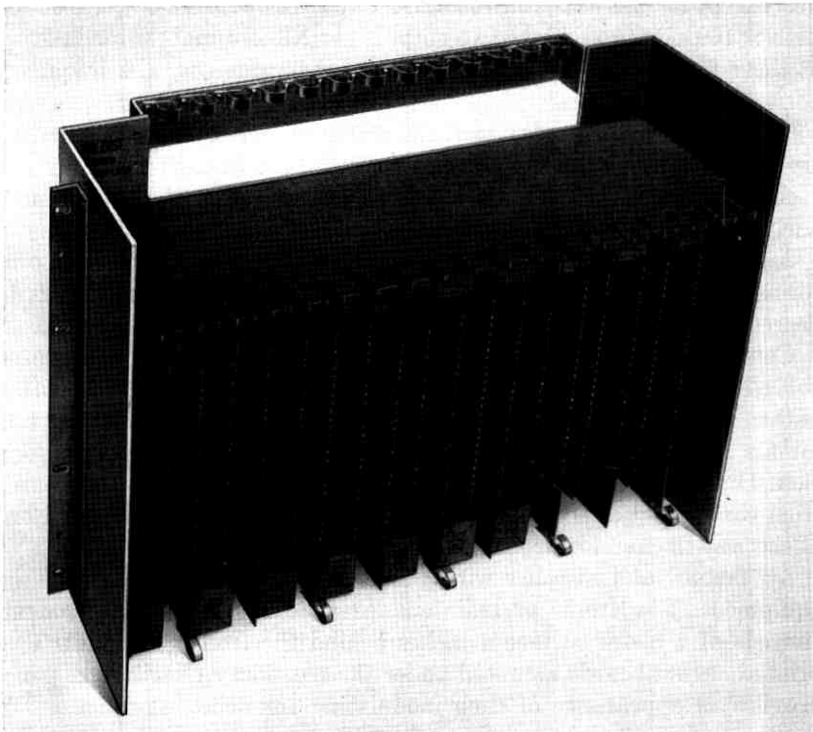

Fig. 26 — Primary distribution panel (front view, cover removed).

of 16 frequencies to each of 13 secondary distribution panels. The second is to provide filtering of the square waves received from the limiting amplifiers, so that pure sine waves tones will be distributed.

The primary distribution panel consists basically of two similar printed wiring board designs. The input wiring to these boards is done at the factory and the distribution from the boards is directly to the installer connected cable. As in the secondary distributing panel, all filtering and distribution is done on the boards and no additional wiring to terminals strips or other distributing means is required.

The smoothing filters are made up of components which cause their height to be greater than that commensurate with the number of distribution boards required in the space available. In order to provide for the filters within the space limitations, two boards were designed which allow the filter on one board to interleave with the adjacent board. A representation of this is shown in Fig. 27.

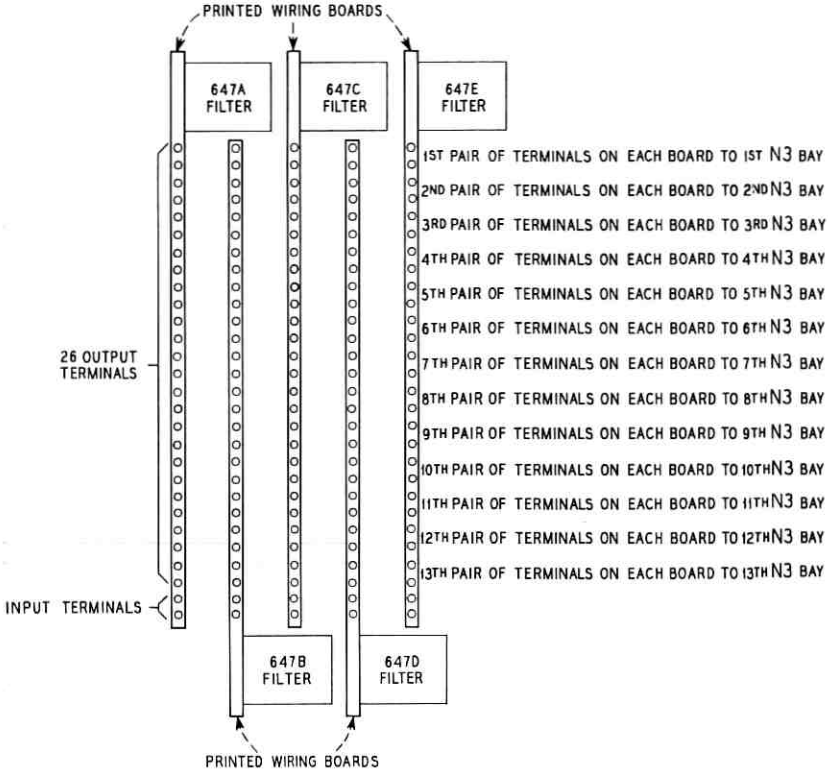

Fig. 27 — Section of primary distribution panel showing mechanical interleaving of smoothing filters and terminal assignment.

Each of the channel frequencies and the N3-L junction frequency is distributed 13 times from an individual board. The balanced low impedance output filter on each board is tuned to the particular frequency that it is to distribute. The series output capacitor of each of the filters is divided into 26 parts to provide for the balanced distribution of up to 13 pairs of wires with both dc and ac isolation. Each output of each printed wiring primary distribution board is provided with a 115-ohm load resistor soldered to the printed wiring path. These resistors terminate unused outputs. The resistors are clipped out and discarded as additional distribution cables are attached. If for some reason one or two of these cables were later disconnected at either end, it would not be necessary to replace the terminating resistors since the isolation between loads is such that an insignificant difference in level would result. No more than two outlets per board should be left unterminated, however.

Connections to these printed boards are made to solderless wrap terminals mounted on the front edge. The major physical difference between the secondary distribution and the primary distribution panels is that the solderless wrap terminals for interconnection are located on the front of the primary distribution panel and on the back of the secondary distribution panel.

The channel group frequencies and the group carrier frequencies are distributed in a manner similar to the distribution of the channel carriers; however, there is no additional fan out of these frequencies in the secondary distribution panel and the primary distribution requires 26 outputs of each frequency. This additional primary distribution is accomplished by providing two printed cards instead of only one for each of these frequencies. A single filter is used for these frequencies, but its output capacitance is divided 52 times to accommodate the 26 outputs. The filters for all of the frequencies are so designed that each of the isolating output capacitors has the same value.

Fig. 28 indicates the input terminals placed at the bottom of one of the distribution cards and the outputs arranged above them. Distribution to a single office cable requires that the installer connect a separate pair of wires from each cable to a pair of terminals on each of the cards. Each double row of terminals across the primary distribution panel provides the required outputs for a terminal bay containing two N3 terminals. This supplies a total of 16 frequencies, three of which are distributed twice through each cable.

The cards and terminals on them are arranged so as to require a minimum number of connections by both the installer and the factory. The terminals are numbered to provide the same relationship found in a terminal block to which the installer normally connects.

The primary distribution panel is provided with a cover to conceal all of the connections since this area faces the maintenance aisle.

A complete view of the carrier supply showing the primary distribution panel, the protecting switching panel, the 4-kc generator and harmonic filters and the plug-in units and switching and alarm panel is shown in Fig. 20. The relationship of the generating and primary distribution portion of the carrier supply to the secondary distribution and the 13 double bays containing 26 N3 terminals is also shown by this figure.

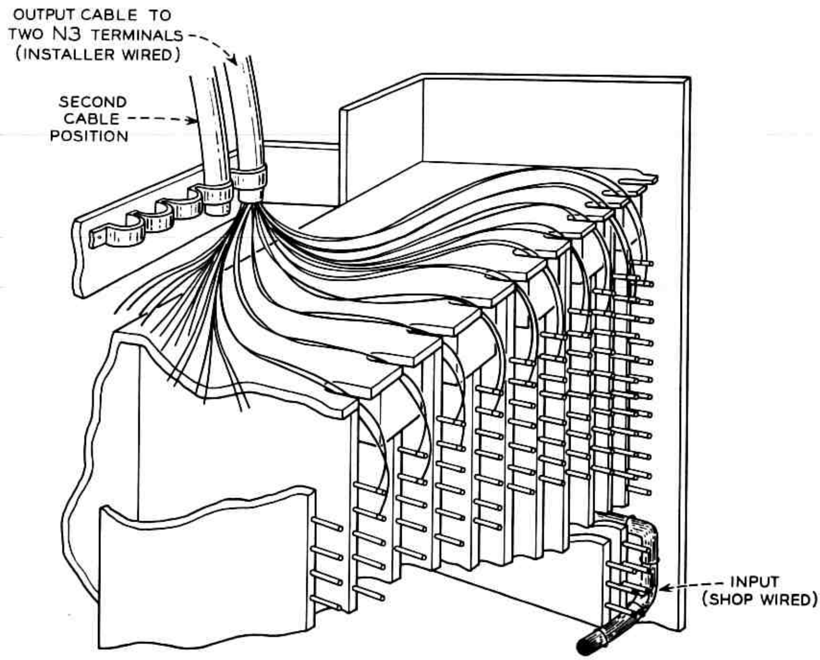

Fig. 28 — Input and output cable connections — primary distribution panel.

XIX. SUMMARY

Equipment design features of the new N3 carrier equipment help substantially in achieving many design objectives. The prime objectives of greatly improved transmission performance, service reliability, and ease of operation and maintenance guided and influenced all equipment design efforts and decisions.

Economy of manufacture results from careful selection of parts and components to assure satisfactory quality at minimum cost. The assembly and wiring of functionally related equipment in a shop-wired package reduces cost, improves noise and crosstalk performance, permits more complete shop testing and simplifies job engineering and installation. Use of the same die cast plug-in unit module frame and the same unit mounting shelf for both terminal and carrier supply achieves economy of manufacturing tooling. Easily removable filters used in certain units reduce the requirements for spare plug-in units.

This new 24-channel terminal provides substantially better performance, requires less space, uses less power, and costs less per installed channel than the equipment it was designed to replace.

REFERENCES

1. Bleisch, G. W. and Irby, C. W., The N3 Carrier System: Objectives and Transmission Features, B.S.T.J., This Issue, pp. 767–799.
2. Haner, R. L. and Wood, I. E., Circuit Design of the N3 Carrier Terminal, B.S.T.J., This Issue, pp. 801–844.

The SNOBOL3 Programming Language

By D. J. FARBER, R. E. GRISWOLD and I. P. POLONSKY

(Manuscript received March 4, 1966)

SNOBOL3 is a programming language designed for the manipulation of strings. Features of the language include symbolic naming of strings and pattern matching. In addition to a basic set of primitive string-valued functions, the system includes the facility for defining functions. These defined functions facilitate the programming of recursive procedures.

This paper presents an intuitive description of SNOBOL3 and at the same time incorporates complete reference material for the programmer. The implementation of SNOBOL3 for the IBM 7094 computer operating under BE-SYS-7 is the basis for this description, but most of the material is common to all implementations.

I. INTRODUCTION

In recent years a number of high-level programming languages have been developed to extend the usefulness of the computer in dealing with primarily nonnumerical problems. The most widely used languages have been IPL,[1] LISP,[2] and COMIT.[3] In 1962 SNOBOL[4] was developed for problems involving the manipulation of character strings. The basic operations of SNOBOL permit the formation, examination, and rearrangement of strings. SNOBOL3 is a generalization and extension of SNOBOL. New features include string-valued functions and input-output facilities integrated into the string structure of the language. There are two types of functions: primitive functions that are included in the system and defined functions that are defined by the programmer in the SNOBOL3 language.

This paper is a description of SNOBOL3 as a programming language. Emphasis is placed on the language as distinct from its implementation. In order to provide information for the potential programmer, however, some references to the implementation are necessary. There are several implementations of SNOBOL3 which differ in detail, particularly with regard to input-output. The implementation for the IBM 7094 computer operating under the BE-SYS-7 monitor is the basis for this paper.

Areas where other implementations are likely to differ are noted in the applicable sections.

Section II describes briefly and informally the essential features of the language. This section is designed as a survey to provide an understanding of the general nature and capabilities of the language. Section III is an elaboration of Section II completing the description of the language. Sections II and III together provide a reference source for the programmer. Section IV describes the environment in which the language operates, including information which the programmer will find useful in running programs. A list of primitive functions and sample programs are included in appendices.

II. INFORMAL DESCRIPTION

SNOBOL3 has just one type of basic data structure: a string of characters. The primitive operations of the language provide for the formation, examination and rearrangement of strings. Arithmetic is defined for operands that are integer strings. The operations to be performed are specified in statements that may also be labeled and may have go-to's specifying transfers. A SNOBOL3 program consists of a sequence of statements terminated by an END statement.

2.1 *Names*

A symbolic name can be assigned to a string and used as a means of referring to that string. There are several ways in which a name can be assigned a value. The simplest is the assignment statement. For example, the statement

```
VOWELS = "AEIOU"
```

assigns the string AEIOU as the value of the name VOWELS. The string consisting of a pair of quotation marks enclosing a string of characters is a *literal* specifying the string AEIOU. The string VOWELS appearing to the left of the equal sign is a *name*. A name that has been assigned a value can be used whenever it is necessary to refer to that value. Thus,

```
NON.CONST = VOWELS
```

causes NON.CONST to have the same value as VOWELS. The null string, having length zero, can be assigned explicitly as value as in the statement

```
ZIP =
```

2.2 *Concatenation*

The basic operation of concatenation of strings is indicated by listing the names successively. The names are separated by blanks. Thus, to concatenate two string names STRING1 and STRING2 and then assign the result to the name STRING3, the following assignment statement suffices:

```
STRING3 = STRING1    STRING2
```

Many strings can be concatenated in a string expression, with literals as well as names used to specify the strings. Thus, the following rules

```
   ARGUMENT = "2X + 3"
EXPRESSION = "SIN("   ARGUMENT   ")"
```

would assign the string SIN(2X + 3) to the name EXPRESSION.

2.3 *Integer Arithmetic*

Arithmetic operations can be performed on integer strings with the operators +, −, /,* having their usual meaning in integer arithmetic and ** indicating exponentiation. Blanks are used to separate the strings and operators. The statements

```
J  = "5"
I  = "3"
N  = I + "2"
M  = (I * "3") + J
```

assign the values 5 and 14 to the names N and M. All arithmetic operations are binary but more complex expressions can be constructed using parentheses as indicated in the last example. Arithmetic has precedence over concatenation, and both types of operations can be performed in one assignment statement. Hence, the statement

```
INDEX = "A." I + "1" "." J
```

assigns the value A.4.5 to the name INDEX.

2.4 *Pattern Matching*

String pattern matching consists of examining a string for a succession of substrings of specified form. A pattern-matching statement consists of the string to be examined followed by a pattern. In its simplest form the pattern may be simply a string. For example, the statement

```
NAME.1    "IS"
```

would examine the value of NAME.1 to determine whether it contains the literal substring IS. The success or failure of a pattern match can affect the flow of the program and has other consequences that will be described later. In the example above, NAME.1, which specifies the string to be examined, is called the *string reference* of the statement. The string reference can also be a literal as in the following pattern-matching statement.

"+−*/" OPERATOR

There are a variety of types of patterns in SNOBOL3 enabling the programmer to make complex inquiries about a string. The pattern, for example, may be expressed as a concatenation of strings as in the statement

EXPRESSION "X" OPERATOR "1"

Patterns of greater generality may be obtained by using *string variables.* As the name indicates, a string variable may have a string as value. There are several types of string variables and the strings which are acceptable values of a string variable depend on the type of the variable. The simplest type of string variable is the arbitrary string variable, so named because it can have any string as its value. An arbitrary string variable is designated by a name bounded by asterisks.

A typical example of the use of an arbitrary string variable would be in determining whether the value of NAME.1 contains the string THE and the string IS but not necessarily consecutively. The arbitrary string variable would be used to match the substring between THE and IS. The pattern-matching statement could be

NAME.1 "THE" *SEPARATOR* "IS"

If the value of NAME.1 were THERAPIST, then the pattern match would be successful with *SEPARATOR* matching RAP.

A consequence of the successful pattern match is the naming of substrings that match string variables. In the above example, SEPARATOR would be given the value RAP as if the assignment statement

SEPARATOR = "RAP"

had been executed.

In addition to arbitrary string variables, there are two other types of string variables: fixed-length and balanced.

A fixed-length string variable can match any string of a specified number of characters. The notation for a fixed-length variable is similar to an arbitrary string variable, except the name is followed by a slash and then by a string specifying the length.

The first three characters of the string named TEXT could be named PART1 by the rule

TEXT   *PART1/"3" *

If N had the value 3, the statement could have been written

TEXT   *PART1/N*

As a second example, consider the statement

"+−*"   *PLUS/"1"*   *MINUS/"1"*   *STAR/"1"*

The pattern successfully matches the string, and PLUS, MINUS, and STAR are assigned values.

A balanced string variable can only match strings that are parenthesis balanced in the usual algebraic sense. Strings matched by balanced string variables do not have to contain parentheses but cannot be null. Such variables are therefore useful for pattern matching on strings that are mathematical expressions. The notation for a balanced string variable consists of a name enclosed within parentheses and surrounded by a pair of asterisks. For example, if EXPRESSION has the value SIN(A*(B + C)), then the pattern match in the statement

EXPRESSION "SIN(" *(ARG)* ")"

is successful and ARG is given the value A*(B + C). This use of the balanced string variable may be compared to the arbitrary string variable in the following example

EXPRESSION "SIN("   *ARG1*   ")"

where the value A*(B + C would be assigned to ARG1.

2.5 *Rearranging Strings*

By combining the operations of scanning and assignment in the same rule, strings may be modified by replacement, deletion, or rearrangement. In particular, if a pattern is followed by an equal sign and then by a string expression, the substring matching the pattern will be replaced by the value of the expression if the pattern match succeeds. As an example of replacement, consider the following sequence of rules.

CARD = "KING OF HEARTS"
CARD   "HEART" = "DIAMOND"

The second statement causes HEART to be replaced by DIAMOND producing the string KING OF DIAMONDS. The following example illustrates how the naming of substrings by string variables may be used in the expression that specifies the rearrangement. The statements

```
SUM = "A1+A2"
SUM    *X* "+" *Y* = "+(" X "," Y ")"
```

change the value of SUM to +(A1,A2).

2.6 *Indirect Referencing*

A level of indirectness can be introduced in SNOBOL3 by prefixing a $ to a name. Thus, if DAY has the value TUESDAY, $DAY is equivalent to TUESDAY. An example of the usefulness of this facility is the ability to modify the naming done in a pattern match. Thus, in the following statements

```
DAY = "TUESDAY"
TEXT     "," *$DAY* ","
```

the name TUESDAY will be assigned to a substring of TEXT if the value of TEXT is such that the pattern match succeeds.

A $ can also be prefixed to a string expression that is enclosed in parentheses. For example, the following statements assign the value of WORD to one of the names LISTA, LISTB, ··· LISTZ according to the first character in the value.

```
WORD *CH/"1"*
$("LIST" CH) = WORD
```

Thus, if WORD has the value DALLAS, the first statement sets the value of CH equal to D. The parenthesized expression

```
("LIST" CH)
```

has the value LISTD. Hence, the effect of the $ is to make the second statement equivalent to

```
LISTD = WORD
```

2.7 *Labels and the Flow of Control*

A label may be assigned to a statement for reference when controlling the flow of the program. The label is merely appended to the beginning of the statement as in

```
HERE  LIST = "(A,B,C,D)"
```

A statement without a label must begin with a blank.

Statements in a SNOBOL3 program are executed in sequential order. This order of execution can be modified by means of a go-to appended to the end of a statement. Go-to's are separated from the rest of the statement by a slash. There are two types of go-to: unconditional and conditional. The unconditional go-to consists of a label enclosed within parentheses. Thus, after executing

```
HERE   LIST = "(A,B,C,D)"            /(THERE)
```

control is transferred to the statement with label THERE. By means of the conditional go-to, control can be transferred depending on whether success or failure has been signaled during the execution of the statement. The letters S and F are used to indicate the two conditions. For example, in the following rule the transfer to the statement with label L2 will occur only if the pattern match is successful.

```
L1      TEXT      "," *A* "."      /S(L2)
```

If the pattern match fails, the next statement in the program is executed.

Transfer on a failure signal can be similarly programmed. As an example, consider the following sequence of statements which will delete from the string named TEXT all occurrences of the characters in LIST:

```
L1     LIST     *CHAR/"1"* =       /F(DONE)
L2     TEXT      CHAR =            /S(L2)F(L1)
DONE
```

In statement L1, the first character in LIST is named CHAR and is deleted by replacing it with a null string. Statement L2 is executed repeatedly until all occurrences of CHAR have been deleted from TEXT. Then the process iterates using the next character in LIST. Finally, when there are no characters left in LIST, the pattern match in statement L1 fails. Thus, if TEXT had the value A+B*C/D+E and LIST the value +*/−, the resulting value of TEXT would be ABCDE.

A transfer may be computed by the use of indirectness in the go-to as illustrated in the following example: If PDL is assumed to have the value A1,B5,C3, then the statement

```
PDL     *RET* "," =      /S($RET)
```

causes deletion of A1, and transfer to the statement labeled A1.

2.8 *Functions*

There are two types of functions in SNOBOL3: primitive and defined. Some functions may signal failure. This failure terminates the execution of the statement in which the function call occurs and may be used to control the flow of the program.

2.8.1 *Primitive Functions*

A basic set of primitive functions, programmed at the machine-language level, has been included in the SNOBOL3 system.

SIZE is an example of a primitive function. The value of SIZE(X) is the number of characters in the string named X. Thus, the statements

```
STR = "FOUL"
  Z = SIZE(STR)
```

assign the value 4 to Z. As a result of the statements

```
X = SIZE(TEXT) - "1"
TEXT     *FRONT/X*     *LAST*
```

LAST is defined to be the name of the last character in TEXT.

One use of functions is to conditionally signal failure and hence alter the flow of control. For example, EQUALS(X,Y) signals failure if X and Y do not have identical values. If the values of X and Y are identical, the function returns the null string as value. Thus, the statement

```
N = EQUALS(A,B) N + "1"
```

will increment N only if the values of A and B are equal.

Another type of primitive function is one that modifies the behavior of the SNOBOL3 system itself. The function call MODE("ANCHOR") is an example. It modifies the pattern matching processor and returns a null value. Subsequently a pattern match can succeed only if the matching substring is at the beginning of the string reference. Thus, if MODE("ANCHOR") has been executed before the statements

```
EXP = "SIN(A + B)"
EXP "(" *(ARG)* ")"
```

the pattern match fails.

2.8.2 *Defined Functions*

A section of SNOBOL3 program can be defined to be a function and certain names occurring in the section can be declared formal parameters. This function declaration is accomplished by a call of the primitive function DEFINE. For example,

```
DEFINE("REVERSE(X)","REV")
```

declares the section of program beginning at the statement with label REV to be a function named REVERSE with a formal parameter X. Suppose REVERSE(X) returns as value the string named X with the characters reversed. Then the portion of program defining REVERSE could be

```
REV    X               *CHAR/"1"*     =      /F(RETURN)
       REVERSE       = CHAR     REVERSE     /(REV)
```

The reserved label RETURN causes return to the place at which the function was called. The name of the function, in this case REVERSE, serves a special purpose. When the function is called, its value is saved and then set to the null string. When transfer to RETURN occurs, its value is the value returned by the function. Thus,

```
Z = REVERSE("ABCDE")
```

assigns the value EDCBA to Z.

2.9 *Statement Format*

A SNOBOL3 statement has a simple format consisting of several fields of arbitrary length separated by blanks. The fields are the label, string reference, pattern, equal sign, replacement expression, and the go-to. A statement may contain some or all of these fields. The replacement statement

```
HERE    TEXT    " "   *WORD*   " " = " "     /S(GOT)
```

has all of the fields.

If the label is omitted the statement must begin with a blank. If the next field is not a go-to, it is considered to be the string reference. Thus,

```
                        /(L7.3)
THERE                   /(HERE)
```

are statements that do not have a string reference. The statement

```
L5     EQUALS(OP,"END")     /S(END)
```

has the string reference EQUALS(OP,"END").

The field following the string reference up to an equal sign or a go-to is the pattern. In a statement without a go-to or equal sign, the pattern is the field following the string reference. Thus, in each of the following statements

```
        STR    A  *B*  "," = B     /S(GREAT)
        STR    A  *B*  "," =
L1      STR    A  *B*  ","         /S(GREAT)
        STR    A  *B*  ","
```

the pattern is

```
A  *B*  ","
```

Note that the elements within the pattern are also separated by blanks.

A statement without a pattern, but containing an equal sign, is an assignment statement. Some examples are:

```
ANS = N + "5"
RES =                 /(READ)
```

The latter example assigns the null string as the value of RES.

No fields are permitted after the go-to field.

III. DETAILED DESCRIPTION

The previous section was an informal description of the basic parts of SNOBOL3. The following section completes this description in a more comprehensive and detailed manner.

3.1 *Names and String Expressions*

3.1.1 *Names*

Names are used to refer to string values symbolically. In addition, names are required for certain parts of statements:

(*i*) string variable names
(*ii*) string references in assignment statements
(*iii*) labels in go-to fields.

Names may be explicit or implicit. Explicit names can consist only of letters, numbers, periods and colons. Examples are:

```
N
STATEMENT.VARIABLE
X:1
37
```

Implicit names, constructed by indirect references, may consist of any nonnull string of characters. Any indirect reference is an implicit name.

For example:

```
$F
$SIZE(N)
$("M" K)
```

Consequently,

```
,,,, = "6"
```

is syntactically *incorrect*, but

```
 INAME = ",,,,"
$INAME = "6"
```

is proper.

The particular characters comprising a name have no significance; a name is merely an identifier. A name may be the same as a label or the name of a function.

3.1.2 *String Expressions*

The basic string-valued elements are:

(*i*) literals
(*ii*) names
(*iii*) function calls
(*iv*) arithmetic operations
(*v*) parenthetical groupings.

Any string of characters (including the null string) not containing a quotation mark (see Section 3.1.3) may be included between the quotation marks of a literal. Function calls, parenthetical groupings, and names may be indirectly referenced. Parentheses are required between successive levels of indirect references.

A string expression is a string-valued element or the concatenation of several such elements. Some typical string expressions are:

```
"PARAGRAPH SUB-HEADINGS FOLLOW"
"N" ((A + "1") * INTERVAL)
$BASE + SIZE(N)
F(X,F(X,X))
$($($ROOT))
M "." P
$("N" I)
```

3.1.3 *Names and Values*

All names have null values at the beginning of program execution except for the string QUOTE. QUOTE has a preassigned value which is a quotation mark.

Names, including QUOTE, may subsequently be given other values by assignment statements or as a result of pattern matching. The resulting name-value relationship between strings forms the basic data structure in SNOBOL3. Structures can be built to arbitrary depths. For example, the statements

```
N1 = "N2"
N3 = "N2"
N2 = "N4"
N4 = "N6"
N5 = "N4"
N6 = "N3"
```

might be used to represent relationships between data as indicated in Fig. 1.

Indirect referencing can be used to refer to the relationships in the structure. The range of such structures is limited by the fact that a name can have at most one value at any time, while a string can be the value of any number of names simultaneously.

3.2 *Arithmetic*

3.2.1 *Integers*

Some strings have the property of being SNOBOL3 integers. Such strings are required in arithmetic operations and as arguments of certain primitive functions. In order for a string to be a SNOBOL3 integer

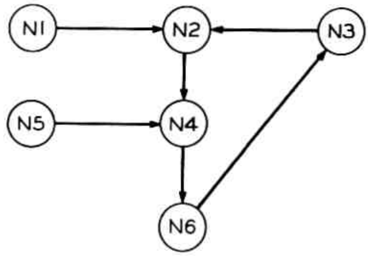

Fig. 1 — The name-value relationship among data.

(*i*) it must consist entirely of digits except for the first character which may be a sign, and
(*ii*) its absolute value considered as a decimal integer must be less than 10^{10}.

In numerical contexts
(*i*) unsigned numbers are taken to be positive,
(*ii*) leading zeros are ignored,
(*iii*) minus zero is equal to plus zero, and
(*iv*) the null string is taken to be zero.

The following strings are SNOBOL3 integers:

```
      5
    +10
0003976
    -37
-000003
     +0
```

The following strings are *not* SNOBOL3 integers:

```
           -
          +A
     3.27E-2
         3.7
876935476271
          0-
      10,000
```

The primitive function .NUM(X) succeeds, returning a null value, if the value of X is a SNOBOL3 integer and fails otherwise. Thus, .NUM("A") fails, while .NUM("100") returns a null value.

3.2.2 *Arithmetic Expressions*

Arithmetic operations must be separated from their operands by blanks. Consequently A+B is syntactically incorrect. Any expression whose value is a SNOBOL3 integer is an acceptable operand.

All arithmetic operations are binary. Thus,

```
N + "3"
```

is a legal operation, while

```
N + "3" * "2"
```

is a syntactic error. Parentheses may be used for grouping terms to create more complicated expressions:

N + ("3" * "2")

In expressions containing both concatenation and arithmetic, arithmetic has precedence over concatenation. Thus, the value of

"N" "5" + "7"

is N12 and the value of

"3" * "2" "10" / "2"

is 65. Parentheses may be used to group concatenations and arithmetic to obtain the desired result. Thus, the value of

"3" * ("2" "10" / "2")

is 75.

The following sequence of statements illustrates possible combinations:

```
ALPHA = "ABCDEFGHIJKLMNOPQRSTUVWXYZ"
    N = SIZE(ALPHA) + "1"
    M = (N + SIZE(N)) * "2"
    K = ("−" N M) + "5"
```

As a result of executing these statements, N would have the value 27, M the value 58, and K the value −2753.

The result of an arithmetic expression is a *normalized* SNOBOL3 integer. Integers are normalized as follows:

(*i*) positive integers are unsigned,
(*ii*) leading zeros are removed, and
(*iii*) any value equal to zero is returned as an unsigned zero.

Thus,

"+0003" + "0"

has value 3, and

"" * "2"

has the value 0.

3.2.3 *The Evaluation of Arithmetic Expressions*

Two modes for evaluating arithmetic expressions are available. The normal mode is *truncation*. In the truncation mode any fractional part resulting from division (or exponentiation) is discarded. Thus, the value of

"5" / "2"

is 2, and the value of

"3" ** "−1"

is 0.

An *integer* mode is available which causes an arithmetic operation to fail if a fractional part would result. The integer mode may be invoked by executing the function call MODE("INTEGER"). The normal mode may be restored by executing MODE("TRUNCATION").

3.2.4 *Error Conditions in Arithmetic Operations*

Error conditions in arithmetic operations occur if:

(*i*) a fractional part would occur in integer mode,
(*ii*) an operand is not a SNOBOL3 integer,
(*iii*) the result of an arithmetic operation is not a SNOBOL3 integer (because it is too large), or
(*iv*) division by zero is attempted.

In all cases, the arithmetic operation fails, terminating the execution of the rule in which it occurs. The failure may be utilized to change the flow of control by means of a conditional go-to.

3.2.5 *Numerical Functions*

There are six functions for comparing the magnitude of integers:

.EQ(X,Y)	$(X = Y)$
.NE(X,Y)	$(X \neq Y)$
.LT(X,Y)	$(X < Y)$
.LE(X,Y)	$(X \leqq Y)$
.GT(X,Y)	$(X > Y)$
.GE(X,Y)	$(X \geqq Y)$.

These functions succeed, returning a null value, if the condition indicated is satisfied and fail otherwise. The functions also fail if either argument is not a SNOBOL3 integer. A common use of the functions is to control loops. For example, the following program assigns the squares of the first 100 positive integers to the names SQ1 through SQ100, respectively.

```
           N = "1"
COMPUTE    $("SQ" N) = N * N
           N = .LT(N,"100") N + "1"   /S(COMPUTE)
```

The function .REMDR(X,Y) has as its value the remainder of X divided by Y. For example, the value of

```
.REMDR("5", "2")
```

is 1. The sign of the remainder is the same as the sign of the divisor and the value is normalized.

.REMDR fails if either argument is not a SNOBOL3 integer or if the value of Y is zero.

3.3 *Pattern Matching*

Pattern matching is a basic operation in SNOBOL3. The examination, rearrangement, and combination of data depend on pattern matching; and the success or failure of matching is often used for altering the flow of control.

3.3.1 *Pattern Elements*

A pattern consists of a succession of pattern elements separated by blanks. There are two basic categories of pattern elements: *string constants* and *string variables.*

Any string expression is a constant, except that arithmetic expressions must be enclosed in parentheses. The following expressions are examples of string constants:

```
K
"35R"
SIZE(Z)
.LE(N,M + SIZE(L))
(M + (N * "2"))
```

String variables may or may not have associated names. The following elements are examples of string variables:

```
**
*()*
*/"3"*
*VARIABLE1*
*$SIZE(N)*
*(EXP)*
```

The length of a fixed-length variable may be any string constant whose value is a nonnegative SNOBOL3 integer when evaluated. The following fixed-length variables illustrate possible forms the length may take:

```
*FL/N*
*V/(SIZE(N) + "1")*
*V/(M + (N * Z))*
```

The lengths of the following variables are syntactically *incorrect:*

HEAD/N + "1"	(Arithmetic expressions must be enclosed in parentheses.)
SPAN/"A"	(The value of the length must be an integer.)

3.3.2 *The Matching Process*

Pattern matching consists of three phases:

(*i*) evaluation of expressions in the pattern,
(*ii*) the actual matching, and
(*iii*) the assignment of values to names associated with string variables.

3.3.2.1 *Evaluation.* Before any matching, all expressions in the pattern are evaluated. Expressions may occur in string constants, the names of string variables, and in the length of fixed-length variables. Evaluation proceeds from left to right. Any failure in evaluation (such as the failure of a function call or arithmetic operation) terminates the execution of the rule without any matching or naming.

The value of all expressions is fixed by evaluation before matching. No evaluation is performed during matching. The only exception to this rule is back referencing described in a following section. Thus, in the pattern

```
*N*     *SPAN/N*
```

the length of the fixed-length variable is the value of N *before* matching and is not influenced by any subsequent match for the arbitrary string variable with the name N.

3.3.2.2 *Matching.* Pattern elements must match consecutive substrings in the value of the string reference. In most cases the match can easily be determined from the following rule:

Pattern matching proceeds from left to right, each pattern element matching the shortest possible substring according to the type of the element.

In some complicated cases, more precise definitions are necessary. The following definitions provide the details for resolving difficult cases.

(*i*) The pattern match proceeds element by element from left to right starting at the leftmost (first) element. The elements must match consecutive substrings in the value of the string reference.

(*ii*) An attempt is first made to match the first element starting at the first character in the value of the string reference. If this is not possible, an attempt is made starting at the second character, and so on.

(*iii*) When an element is successfully matched, a *forward* match is attempted for the next element.

(*iv*) If an element cannot be matched, *rematch* is attempted for the preceding element. Rematching is an attempt to extend the substring matching a pattern element and occurs because the pattern match cannot be successfully concluded with the previous match.

(*v*) Pattern matching terminates successfully when the rightmost (last) pattern element has been matched. Pattern matching terminates in failure if no match can be found for the first element.

The methods of forward matching and rematching depend on the type of the pattern element. In each case, the element must match a substring in the string reference starting at the character following the substring matching the preceding element. The details follow.

(a) *String Constants*

In forward matching, a string constant matches a substring identical to its value. If this is not possible, forward matching fails. A null constant always matches.

No rematch is possible, and rematching always fails.

See the special case of back referencing.

(b) *Arbitrary String Variables*

In forward matching, an arbitrary variable matches a null string.

In rematching, one character is added to the substring previously matched by the variable. If the string reference is not long enough for such a match, rematching fails.

As a special case, if the last element in the pattern is an arbitrary string variable, it matches the remainder of the string.

(c) *Balanced String Variables*

In forward matching, the string matched by a balanced variable depends on the first character of the substring where the variable is to match. If this first character is not a parenthesis, then the variable matches that character. If the first character is a right parenthesis, the match fails. If the first character is a left parenthesis, the string being examined is considered character by character until a matching right parenthesis is found. If there is no matching parenthesis, failure is indicated. Notice that a balanced string variable always matches at least one character.

In rematching, the previously matched substring is extended by the next shortest balanced string according to the rules for forward matching. If this is not possible, rematching fails.

(d) *Fixed-Length String Variables*

In forward matching, a fixed-length variable matches a substring of length specified by the variable. If the string being examined is not long enough, forward matching fails.

Rematching always fails.

(e) *Back Referencing*

Back referencing is a special case in pattern matching in which tentatively matched substrings can be referred to dynamically during the matching process. If a constant in the pattern has the same name as a name associated with a variable to the left of it in the pattern, the value of the constant is taken to be the substring currently matched by the variable. Thus, in the pattern

```
*N* "," N
```

the constant N must match a substring identical to the substring matching *N*. Since matching is done left to right, a tentative match always exists for a back-referenced variable.

Back referencing only occurs when the name associated with a variable is to the left of a constant with the same name. Consequently the pattern

```
N "," *N*
```

does not contain back referencing.

If there are several occurrences of the same name in a pattern, a named constant back references the variable with its name which is nearest to it on the left. In the pattern

```
*N* "," N N *N* "," N
```

the first and second named constants refer to the first variable and the third named constant refers to the second variable.

Any type of variable may be back referenced and any number of named constants may back reference variables in an arbitrarily complicated way.

The determination of back referencing within a pattern is made after the evaluation of expressions in the pattern but before matching. In the statements

```
A = "C"
B *C* "," $A
```

the pattern is back referenced. However, in the statements

```
VARI = "SPAN"
X  *VARI*  $VARI
```

there is no back referencing.

Back referencing only applies to names which are pattern elements and not to any other name in the pattern. Specifically in the pattern

```
*N*  *INT/SIZE(N)*
```

the length of INT is determined by evaluation before matching and does not change during the matching process.

3.3.2.3 *Naming.* If the pattern match fails, no naming is done and the execution of the rule is terminated. If the pattern match succeeds, naming is performed from left to right for each name associated with a string variable. The substring matching the variable becomes the new value of the associated name. If a name is associated with more than one variable, the value is assigned corresponding to the rightmost variable with that associated name.

In the case that a name is computed as the result of an expression, the name is determined by the evaluation made before pattern matching. Thus, in the statements

```
A = "C"
Z  *A*  ","  *$A*
```

the name associated with the second string variable is C regardless of the value of Z.

3.3.3 *Pattern Matching Modes*

In the normal mode of pattern matching, the first element of the pattern may match starting anywhere in the value of the string reference. Thus, the simple match

"0123456789" "6"

succeeds. This mode is referred to as *unanchored*. The alternative mode, in which the first pattern element must match a substring beginning with the first character of the string reference, is called *anchored*. This mode may be invoked by executing the function call

MODE("ANCHOR")

Subsequently, all pattern matching will be in the anchored mode unless otherwise modified. The normal mode may be restored by

MODE("UNANCHOR")

The mode of matching may be changed for the duration of a single statement by means of the two functions ANCHOR and UNANCH. These functions, which have no arguments, must be called before matching (Refer to the Section 3.4.3). Both functions return null values. Thus, in

Z ANCHOR( ) "." *IDENT* "."

the pattern match is anchored regardless of the matching mode current in the program. Subsequent statements are not affected. The null value returned by ANCHOR does not affect the match since a null string may match anywhere.

ANCHOR and UNANCH supersede the MODE function even if the MODE function is executed subsequently in the evaluation of the pattern elements. Hence in the statement

STRING UNANCH( ) MODE("ANCHOR") "."

the pattern match is unanchored, although the anchored mode will subsequently prevail.

3.3.4 *Examples of Pattern Matching*

The following examples illustrate some of the situations which occur in pattern matching. String reference values are given as literals for clarity. Naming is indicated for those pattern matches which succeed. The normal unanchored matching mode is assumed.

Example 1:

"K)AK(A + B + C)ST" "K" *(A)* "ST"

The match succeeds with

A = "(A + B + C)"

Example 2:

"K)AK(A + B + C)ST" ANCHOR( ) "K" *(A)* "ST"

The match fails.

Example 3:

"S)(S + A*B(S" "S" *(A)* "S"

The match fails.

Example 4:

"ABCDEFGHIJKLMNO" *HV/"5"* *A* "K" *B*

The match succeeds with

HV = "ABCDE"
A = "FGHIJ"
B = "LMNO"

Notice that since the last pattern element is an arbitrary string variable it matches the remainder of the string reference.

Example 5:

"364:" *A* *SUM/"3"* ":"

The match succeeds with

A = " "
SUM = "364"

Example 6:

"ARMY" *A* *B* *C*

The match succeeds with

A = " "
B = " "
C = "ARMY"

Notice that the first two arbitrary string variables match null strings since this satisfies the requirement for matching the shortest possible substrings.

Example 7:

"ABC" *(BAL1)* *(BAL2)*

The match succeeds with

BAL1 = "A"
BAL2 = "B"

Example 8:

"AB" *(BAL1)* *(BAL2)* *(BAL3)*

The match fails since each balanced string variable must match at least one character.

Example 9:

"ABCD" *S/"2"* *T/"3"*

The match fails since the string being matched is not long enough.

Example 10:

"ABCDEFGHFGH" *A/"3"* A

The match succeeds with

A = "FGH"

This is a simple example of back referencing.

Example 11:

"ABCDEFGHFGH" ANCHOR( ) *A/"3"* A

The match fails.

Example 12:

"32579.97" *A* *B* "." B A

The match succeeds with

A = "7"
B = "9"

These values can be verified by carefully applying the matching rules. (The expected match might be a null value for both A and B.)

Example 13:

The following example illustrates the complexity which may occur with back referencing.

"BACCABACABABACACAB" *A* *(B)* *(C)*
D C D B D C A *E* A E

The match succeeds with

A = " "
B = "BAC"
C = "CAB"
D = "A"
E = " "

Example 14:

"A,A,B,B" *X* "," X "," *X* "," X

The match succeeds with

X = "B"

3.4 *Program Structure and the Flow of Control*

A program consists of a succession of statements terminated by an END statement containing the reserved label END. The END statement may also contain a name which is the label of the first statement to be executed. If the END statement contains no name, execution begins with the first statement of the program.

Statements are subsequently executed one after another unless control is transferred by means of a go-to.

3.4.1 *Labels*

Labels are distinguished by beginning in Column 1. A statement with no label must have a blank in Column 1. The first character of a label must be a letter or a digit. Subsequent characters may be anything but blanks. Labels are program constants; the particular characters in a label have no significance even if they resemble some other structure such as a name or a function call. Thus F(X) is a legitimate label but has no further meaning.

3.4.2 *The Go-To Field*

Go-to's are used to alter the ordinary sequential execution of statements. In general, a statement may be successfully completed, or failure may be indicated for a number of causes. The success or failure may be sensed and used by corresponding conditional go-to's to alter the order in which statements are executed.

A statement with an unconditional go-to may not have conditional go-to's. Furthermore, a statement may not have more than one unconditional, success or failure go-to. In statements with both success and failure go-to's, the go-to's may occur in either order.

The labels given in the go-to's must be names and transfer is made to the name (not its value). The label in a go-to may be computed by the use of implicit names resulting from indirect references. For example, in the statement

```
X = "3"      /($("R" X))
```

transfer is made to the statement with label R3. Function calls occurring in go-to's must not fail.

3.4.3 *The Order of Execution Within a Statement*

The order of execution of operations within a statement may be important to the programmer for two reasons:

(*i*) Failure of an operation within a statement terminates execution of the statement at that point so that subsequent operations are not performed.

(*ii*) Calls of defined functions may change the values of names which appear subsequently in the same statement.

Consequently, a detailed knowledge of when various parts of a statement are evaluated may be required to determine how a program will function. The overall order of execution within a statement is as follows:

(*i*) The string reference (if any) is evaluated.

(*ii*) The elements of the pattern (if any) are evaluated from left to right (see Section 3.3.2).

(*iii*) The pattern match (if any) is performed.

(*iv*) Any naming as the result of a successful pattern match is performed.

(*v*) If a string expression is specified as a replacement, that string expression is evaluated.

(*vi*) Reformation (if specified) of the value of the string reference is made.

(*vii*) The go-to (if any) corresponding to the success or failure of the statement is evaluated.

(*viii*) Transfer is made to the next statement accordingly.

If failure is signaled in any of the steps (*i*) through (*vi*) above, execution of the statement terminates at that point and the appropriate go-to is evaluated. In particular note that *only* the appropriate go-to is evaluated. The order of evaluation within a string expression is as follows:

(*i*) Elements in a concatenation are evaluated left to right.

(*ii*) In a function or parenthetical grouping, the innermost expression in the nesting is evaluated first.

(*iii*) Arithmetic is performed before concatenation.

(*iv*) All arguments of a function are evaluated, left to right, before the function is called.

3.4.4 *Termination of Execution*

Program execution is usually terminated by a transfer to the label END or by flowing into the END statement.

Depending on the monitor system under which SNOBOL3 operates, the termination of a SNOBOL3 program may or may not terminate the job which initiated the SNOBOL3 program. Thus, two modes may be distinguished:

(*i*) *endjob*, in which job execution is terminated upon completion of the SNOBOL3 program, and

(*ii*) *system*, in which job execution may continue after completion of the SNOBOL3 program.

The normal mode in SNOBOL3 is endjob. The alternative mode may be invoked by the function call MODE("SYSTEM"). The normal mode may be restored by the function call MODE("ENDJOB")

Execution of a SNOBOL3 program may be interrupted by the function SYSTEM(FILE). A call of SYSTEM(FILE) suspends program execution, returning control to the monitor. The input source for the monitor is set to the value of FILE. If the value of FILE is null, the input source is set to the standard input source. The availability of SYSTEM depends on the monitor under which SNOBOL3 operates.

3.5 *Input-Output and File Manipulation*

3.5.1 *Implementation Differences*

Input-output is particularly subject to differences in machines and monitor systems. Consequently the input-output behavior of the

SNOBOL3 system may vary considerably in different implementations. Reference to files, record sizes and the handling of end-of-file differ most. The following sections should be read with this in mind.

3.5.2 *String-Oriented Input and Output*

Input and output is accomplished through string names associated with logical files.

For example, SYSPOT ("system peripheral output tape") is associated with the standard output file. Every time SYSPOT is given a value, a copy of the value is printed on the system output file. Thus, the statement

SYSPOT = "TABLE OF VALUES"

will cause the printing of TABLE OF VALUES on the output listing.

SYSPPT ("system peripheral punch tape") is associated with the standard punch file. Values given SYSPPT are punched rather than printed.

Similarly, SYSPIT ("system peripheral input tape") is associated with the standard input file. Every time the value of SYSPIT is required, a card image is read from the input file to become the value of SYSPIT. For example, the statement

SYSPIT *FIELD1* "," *FIELD2* " "

might be used to read and name data items on input cards with the format indicated by the pattern.

3.5.3 *The Association of String Names with Files*

The names SYSPOT, SYSPPT, and SYSPIT are automatically associated with standard files at the beginning of program execution. Any name (including these three) may be associated with any file during program execution by means of association functions. All the following functions return null values.

(*i*) PRINT(NAME,FILE) associates the value of NAME with the value of FILE in the print sense. Thus,

PRINT("X", "OUT")

associates the name X with logical file OUT. After execution of this function call, copies of all values assigned to X will be placed on the file OUT.

(*ii*) PUNCH(NAME,FILE) associates the value of NAME with the value of FILE in the punch sense. The distinction between punch and print association is described in the next section.

(*iii*) READ(NAME,FILE) associates the value of NAME with the value of FILE in the read sense.

The execution of an association function detaches the name from any file with which it may be associated. Thus,

PUNCH("SYSPOT", "SCR")

is permissible.

Any nonnull name may be associated with any file. If the value of FILE is null, the name will be associated with the appropriate standard file. A name may be associated with only one file, but any number of names may be associated with a file.

3.5.4 *Output*

Output occurs whenever an output-associated name is given a value. Thus,

```
FIELD    *SYSPOT*    ","    *SYSPOT/"5"*
```

results in two outputs if the pattern match is successful.

Print and punch association differ in carriage control. When output is performed on values whose names are associated in the print sense, six blanks are prefixed to provide carriage control. No carriage control is provided for output resulting from punch association. Consequently, punch should be used for intermediate files which are to be subsequently read.

Printing on the standard print file is 126 characters per line (not counting carriage control). Additional lines are generated as necessary for longer strings. Punching on the standard punch file is 72 characters per card with additional cards generated as necessary. Values punched or printed on other files are augmented with blanks to an even multiple of six characters. Resulting strings containing 84 characters or less are printed as single records of the length of the augmented string. Longer records are printed 84 characters per record. Any residual string over a multiple of 84 characters is printed as a record of the residual length.

Printing a null value always produces a record because of the six blanks prefixed for carriage control. Punching a null value does not produce a record.

Since output of strings may break one string into many records, care must be taken that the strings may be properly reconstructed if necessary. This remark also applies to padding with blanks and handling of null strings as described above.

Names associated with output files retain their values like any other names. The output process does not destroy values.

3.5.5 *Input*

All strings read from the standard input file are 84 characters long. Blanks are used to fill out shorter records. Records read from other files are not extended.

It is particularly important to notice that any use of the value of a read-associated name results in the reading of a record and the loss of the previous value of the name. This is true regardless of context. For example,

```
SYSPIT    *Z*    ","
```

results in the reading of a record. The record is destroyed by any subsequent use of SYSPIT. Consequently, it is good programming practice to assign the result of reading to some other string. For example,

```
SYSPOT = SYSPIT
```

prints the next record, and the resulting string remains as the value of SYSPOT for further use.

An important aspect of a read-associated name is the indication of failure if the read operation fails (as the result of an end-of-file or binary record). Such a failure terminates execution of the statement in which it occurs and may be used by a conditional go-to. The ability of the SNOBOL3 system to regain control after an end-of-file depends on the monitor under which SNOBOL3 operates.

3.5.6 *Other Functions*

Several functions exist for performing standard input-output operations and file manipulation. All these functions return null values.

(*i*) DETACH(NAME) removes the value of NAME from any input-out association. For example,

```
DETACH("SYSPOT")
```

terminates normal print output. If the value of NAME is not associated with a file, no action is taken.

DETACH may be used to save the value of a name associated in the read sense. For example, the statements

```
SYSPIT     *A*     ","     *B*     ","     *C*     /S(PROG)
DETACH("SYSPIT")     /(ERROR)
```

might be used to go to an error routine in case an input record does not have the expected format. By detaching SYSPIT, the record in error may be examined without destroying it.

(*ii*) REWIND(NAME) rewinds the file associated with the value of NAME. Note that the argument is the name and not the file. An end-of-file is written on a file in output status before it is rewound.

(*iii*) BSREC(NAME) backspaces one record on the file associated with the value of NAME.

(*iv*) EJECT(NAME) writes an eject carriage control character on the file associated with the value of NAME.

(*v*) OPEN(KEY,FILE) opens the specified file in the key area which is the value of KEY. The applicability of this function depends on the monitor under which SNOBOL3 operates.

3.6 *Primitive Functions*

3.6.1 *Function Calls*

Function calls may occur anywhere in a statement where a string value is appropriate. An argument of a function may be any string expression, however complicated. Any argument may be explicitly null and trailing arguments that are omitted are given null values. Thus, the three function calls

```
EQUALS(X," ")
EQUALS(X,)
EQUALS(X)
```

are equivalent. A primitive function may be called with up to six arguments, regardless of the number specified by the function. Such additional arguments are evaluated but are ignored by the function.

All function calls return strings as value if they succeed. In the case of functions that have no natural value, a null string is returned.

It is important to notice that the function name and the left parenthesis may not be separated by blanks. Thus,

```
SIZE(N)
```

is a function call, while

```
SIZE   (N)
```

is the concatenation of a name and a parenthetical grouping. Similarly, functions such as ANCHOR( ) which have no argument must be written with the parentheses. Otherwise they will be taken for string names rather than function calls.

3.6.2 *Functions Relating to Functions*

(*i*) OPSYN(NEW,OLD) OPSYN permits the programmer to associate a new name with a function. Thus, for example,

OPSYN("LENGTH","SIZE")

makes the name LENGTH a function with the same definition as SIZE. Either LENGTH or SIZE may now be used to call the function. New names may be associated with either primitive or defined functions. OPSYN returns a null value.

(*ii*) CALL(FNC) CALL permits the programmer to invoke a function implicitly by interpreting a string as a function call. The value of FNC must correspond syntactically to a function call.

For example, the function call

CALL("SIZE(M)")

is equivalent to the function call

SIZE(M).

The arguments in FNC are interpreted as explicit names. Thus, all arguments in a function invoked by CALL must be assigned names. For example, to invoke

.GT(SIZE(N),"5")

by the use of CALL, statements of the form

```
ARG1 = SIZE(N)
ARG2 = "5"
CALL(".GT(ARG1,ARG2)")
```

are required.

Any primitive or defined function may be invoked by CALL. Value is returned and success or failure indicated in the same manner as if the function call appeared explicitly.

3.6.3 *Miscellaneous Primitive Functions*

There are six primitive functions in addition to the functions described elsewhere in Section III. They are:

(*i*) EQUALS(X,Y) EQUALS returns a null value if the value of X is identical to the value of Y and fails otherwise. The values must be identical and not just numerically equal (compare Section 3.2.5).

(*ii*) UNEQL(X,Y) UNEQL returns a null value if the value of X is not identical to the value of Y and fails otherwise.

(*iii*) TRIM(S) TRIM returns as value the value of S with trailing blanks removed, thus, for example, TRIM(SYSPIT) is a convenient method of removing superfluous blanks from input cards.

(*iv*) TIME( ) TIME is a function of no arguments which returns as value the millisecond time from the beginning of program compilation. The value is returned as a 6-character number.

(*v*) DATE( ) DATE is a function of no arguments which returns as value the current date. The value is returned as a 6-character number. For example, April 1, 1966 would appear as

040166

(*vi*) SIZE(S) SIZE returns as value the number of characters in the value of S. For example, the value of SIZE ("0123456789") is 10.

3.6.4 *Addition of Primitive Functions to the SNOBOL3 System*

The SNOBOL3 system is designed so that separately-compiled primitive functions may be added easily. This facility has been used extensively to add a wide range of capabilities.[5,6,7] A discussion of the format and communication conventions required for primitive functions is beyond the scope of this paper. Ref. 8, which is available from the authors, describes these matters in detail.

3.7 *Defined Functions*

3.7.1 *The Definition of a Function*

A defined function is characterized by four items:

(*i*) a name, by which it is called and which is used for returning value,

(*ii*) a list of formal arguments, used for passing values to the function,

(*iii*) a label, indicating its entry point, and

(*iv*) a list of local names used by the function.

A function must be defined during program execution before it is used. This definition is accomplished by a call of the DEFINE function which establishes the four items above. The form of the call is

DEFINE(FORM,LABEL,NAMES)

FORM is a prototype of the function call, giving the function name and the list of formal arguments. The value of LABEL is the entry point, and the value of NAMES is the list of local names separated by commas. For example,

DEFINE("FACT(N)","F")

defines a function FACT with one formal argument N. Execution of FACT is to begin at the label F. No local names are declared. Similarly,

DEFINE("MATRIXADD(A,B)","MA","I,J,K")

defines a function MATRIXADD with the two formal arguments A and B, the entry point MA and local names I, J, and K.

The total number of formal arguments and local names must not exceed ten. This limit is an assembly parameter.

3.7.2 *The Execution of Defined Functions*

The call of a defined function is identical to the call of a primitive function (see Section 3.6.1). Hence, trailing arguments which are omitted are given null values. A defined function, however, may not be called with more arguments than given in its definition.

When a defined function is called, values of the following names are saved:

(*i*) the name of the function.
(*ii*) all formal arguments.
(*iii*) all local names.

New values are assigned to these names as follows:

(*i*) the name of the function is given a null value.
(*ii*) the formal arguments are assigned values by evaluating the corresponding arguments in the function call.
(*iii*) all local names are assigned null values.

Saving of old values and assignment of new values is made from left to right as the names appear in the DEFINE call.

After these new values have been assigned, control is transferred to the entry point of the function and program execution continues in a normal fashion until transfer is made to one of the two reserved labels RETURN or FRETURN.

RETURN terminates execution of the function. By convention, the value of the function call is the value of the function name when the return was made. For example, if FACT as defined above is designed to compute the factorial of a number, the corresponding program might be

```
F  FACT = .EQ(N,"0") "1"   /S(RETURN)
   FACT = FACT(N − "1") * N   /(RETURN)
```

Then the statements

```
SYSPOT = FACT("3")
SYSPOT = FACT("2") + FACT("4")
```

would print 6 and 26, respectively.

FRETURN terminates execution of the function and signals failure. Execution of the statement in which the call occurs terminates at that point in the same manner as in the failure of a primitive function.

When return is made from a function (by either RETURN or FRETURN) the saved values of all names are restored in the opposite order from which they were saved.

A function may have a formal argument which is the same as its name. This is useful when the value of a function is to be a simple modification of one of its arguments. A function whose value is its first argument with all occurrences of its second argument deleted might be defined as

```
DEFINE("DELETE(DELETE,CHAR)","DEL")
```

with the program

```
DEL  DELETE  CHAR =  /S(DEL)F(RETURN)
```

Here DELETE can be operated on as desired and the value has the correct name (that is, the name of the function) when the deletion is completed.

3.7.3 *Local Names*

Local names may be declared when names used in a function have values which should not be destroyed by a function call. Consider the following function which intersperses commas between the characters in its argument

```
COMMA ARG    *CHAR/"1"* =  /F(RETURN)
        COMMA = COMMA CHAR ","  /(COMMA)
```

The definition would be

```
DEFINE ("COMMA(ARG)","COMMA","CHAR")
```

so that the use of CHAR during the function call would not change the value of CHAR outside the function.

Local names are particularly important when recursively called functions use names for intermediate computation. Appendix II contains a program in which such use of local names is necessary.

IV. OPERATING ENVIRONMENT

The SNOBOL3 system consists of a compiler and an interpreter. The compiler translates SNOBOL3 source programs into an internal language suitable for the interpreter.

4.1 *Compilation*

4.1.1 *Source Program Listing*

During compilation, the source program is read and compiled card by card. Only columns 1 through 72 are read by the compiler. Consecutive statement numbers are added to the listing for reference.

4.1.2 *Comments*

A card with an asterisk in column 1 is treated as a comment. Comments are printed but otherwise ignored by the compiler. Comments may be used freely throughout the program and may be placed anywhere before the END card.

4.1.3 *Continue Cards*

A statement may be broken over card boundaries by use of the continue card convention. A period in column 1 is interpreted by the compiler as an indication that the card is a continuation of the preceding statement. Statements may be broken over card boundaries anywhere a blank is permissible in the syntax. Literals cannot be broken over card boundaries. A very long literal must be represented as a concatenation of shorter literals. For example,

```
        SYSPOT = "THE MAXIMUM LENGTH OF "
.       "THE COMPUTATIONAL THREAD HAS BEEN "
.       "COMPUTED TO BE"
```

There is no limit to the number of continue cards which may be used for a statement.

4.1.4 *Compiler Control Cards*

The programmer can perform some operations during the compilation process by means of compiler control cards. Compiler control cards are identified by a minus sign in column 1. The control action is taken when the card is encountered. Following the minus sign, the first nonblank subfield is taken to be the control word for the card. Other subfields, if any, are separated internally by commas. The control cards and actions are:

(*i*) — TITLE
Eject to a new page in the listing of the source program. Title subsequent pages with the information on the remainder of the control card.

(*ii*) — EJECT
Eject to a new page in the listing of the source program.

(*iii*) — SPACE
Print a blank line in the source program listing.

(*iv*) — UNLIST
Stop listing the source program. (The source program is normally listed.)

(*v*) — LIST
Resume listing.

(*vi*) — PCC
Print control cards. (Control cards are normally not printed.) PCC is a binary switch. Successive uses turn printing of control cards on and off.

(*vii*) — OPEN KEY,FILE
Open the file in the specified key area.

(*viii*) — REWIND FILE
Rewind the specified file.

(*ix*) — SOURCE FILE
Change the input source for the SNOBOL compiler to the specified file.

(*x*) — SYSTEM
Return control to the monitor under which SNOBOL3 operates.

(*xi*) — NULLOP OP
Make the control card operation OP inoperative for the rest of the program.

Invalid control cards, i.e., control cards not in the list above or with a format error, are printed but otherwise ignored. Control cards may be used anywhere in the program before the END card, including between continue cards.

4.1.5 *Diagnostic Messages from the Compiler*

At the end of compilation one of three comments appears:

(*i*) SUCCESSFUL COMPILATION, indicating the source program contains no syntactic errors,
(*ii*) ERROR IN COMPILATION, indicating syntactic errors in the source program, or
(*iii*) FATAL ERROR ENCOUNTERED DURING COMPILATION, indicating the occurrence of an error of sufficient severity to terminate the compilation process.

Syntactic errors are printed following the source deck listing with statement numbers referring to each type of syntactic error. Compilation of a statement ceases when a syntactic error is encountered. Consequently subsequent errors in the same statement will not be detected. The syntactic error messages are

(*i*) ILLEGAL CONSTRUCTION, usually indicating an illegal character in a name, arithmetic operators not surrounded by blanks, or arithmetic operations in the pattern not enclosed in parentheses.
(*ii*) ERROR IN GROUPING, usually indicating unbalanced parenthesization, e.g., (A B))
(*iii*) TOO MANY ELEMENTS IN FUNCTION OR GROUPING, indicating overflow of an internal buffer due to an excessively complicated parenthetical grouping or function call. The maximum number of elements in such structures is an assembly parameter of the SNOBOL3 system and is about 50.
(*iv*) VARIABLE WITH GROUPING OR FUNCTION NOT CLOSED, indicating that a parenthetical grouping or function call in a string variable is not followed by a terminating asterisk, e.g., *T/SIZE(N)
(*v*) ERROR IN LENGTH SPECIFIER, indicating a syntactic error in the length of a fixed-length string variable, e.g., *F/"A"*
(*vi*) "NAMELESS" STRING REFERENCE IN ASSIGNMENT STATEMENT, indicating an attempt to assign a value to a literal, function call, or parenthetical grouping, e.g., "3" = "2"

(*vii*) ARITHMETIC OPERATION WITHOUT SECOND OPERAND, e.g., A = B + /(L1)
(*viii*) ARITHMETIC OPERATION WITHOUT FIRST OPERAND, e.g., A = + B
(*ix*) TWO ARITHMETIC OPERATIONS IN A ROW, e.g., A = B * * C
(*x*) NONBINARY ARITHMETIC OPERATION, e.g., A = B * C + D
(*xi*) PRIOR STATEMENT NOT PROPERLY TERMINATED, indicating a missing continue card or incomplete construction. This error is not detected until the following card is read and determined not to be a continue card. Consequently a statement such as A = B + not followed by a continue card will cause this error message. Compare with (*vii*) above.
(*xii*) ERROR IN GO-TO FIELD, e.g., /S(L1)(L2)
(*xiii*) "NAMELESS" STRING VARIABLE, e.g., *SIZE(SYSPOT)*
(*xiv*) ILLEGAL LABEL, indicating a label which does not start with a number or letter.
(*xv*) MULTIDEFINED LABEL, indicating the same label has occurred more than once.
(*xvi*) CONTINUE CARD NOT PRECEDED BY STATEMENT, indicating the first card in the source deck is a continue card. (Comment cards and compiler control cards may be freely interspersed between continue cards.)

If a fatal error occurs during compilation, the nature of the fatal error is printed followed by a listing of any syntactic errors. The fatal error messages are

(*i*) PROGRAM BUFFER OVERFLOW, indicating the source program is too large for an internal buffer. The size of this buffer is an assembly parameter.
(*ii*) ERROR READING INPUT TAPE, indicating a binary record was encountered during compilation. This condition is almost always due to the omission of an END statement.
(*iii*) END TRANSFER ADDRESS IN ERROR, indicating the label specified on the END card does not start with a number or letter.
(*iv*) END TRANSFER SPECIFIES UNDEFINED LABEL, indicating the label specified on the END card does not occur as a label in the program.

(*v*) MORE THAN 50 NONFATAL ERRORS, indicating the occurrence of more than 50 syntactic errors in the source program. This limit on syntactic errors is an assembly parameter. Such an excess of syntactic errors usually indicates the source deck is not a SNOBOL3 program.

(*vi*) SYSTEM ERROR, indicating a programming error in the SNOBOL3 compiler, or a machine error.

4.2 *Program Execution*

If a fatal error does not occur during compilation, execution is entered. Execution continues until the program transfers to or flows into the END statement or until an error occurs.

4.2.1 *Error Diagnostics*

The possible errors are:

(*i*) ATTEMPT TO EXECUTE STATEMENT WITH COMPILATION ERROR. Execution is terminated if an attempt is made to execute a statement with a compilation error. In the case of multidefined labels, the first occurrence is considered the valid label, and all transfers to this label go to the first occurrence. Subsequent statements with the same label are considered to be erroneous and flowing into such a statement terminates execution.

(*ii*) ATTEMPT TO TRANSFER TO AN UNDEFINED LABEL.

(*iii*) STRING OVERFLOWED 5461 CHARACTERS. This maximum length of strings is an implementation constraint.

(*iv*) INTERNAL BUFFER OVERFLOW, indicating an internal buffer has been exceeded, usually the result of excessive depth of recursive function calls or an excessively long pattern. The buffer sizes are assembly parameters.

(*v*) ATTEMPT TO USE NEGATIVE LENGTH IN A VARIABLE, indicating the length of a fixed-length variable is negative.

(*vi*) FUNCTION FAILED IN GO-TO FIELD, indicating a function call failed while evaluating a go-to field.

(*vii*) ATTEMPT TO ASSOCIATE A NULL NAME WITH I/O FILE.

(*viii*) ILLEGAL FILE OR FILE OPERATION, such as performing an input or output operation on a name not associated with a file or attempting to rewind the standard input file.

(*ix*) ATTEMPT TO READ PAST EOF ON SYSTEM INPUT TAPE. The first attempt to read an end-of-file on the standard input file results in failure of the statement in which the attempt occurred. A second attempt is fatal.

(*x*) IMPROPER ATTEMPT TO OPSYN A FUNCTION, indicating an attempt to OPSYN a name to an undefined function.

(*xi*) ATTEMPT TO CALL AN UNDEFINED FUNCTION.

(*xii*) IMPROPER DEFINITION OF A FUNCTION, indicating an error in a call of DEFINE.

(*xiii*) UNDEFINED OR NULL LABEL USED IN DEFINE STATEMENT, indicating the label specified in a call of DEFINE is null or does not occur in the program.

(*xiv*) TOO MANY ARGUMENTS IN A FUNCTION DEFINITION, indicating that the number of arguments and local variables in a defined function exceeds ten.

(*xv*) IMPROPER CALL OF A DEFINED FUNCTION, indicating too many arguments in the call of a defined function, or an improper argument for the CALL function.

(*xvi*) FUNCTION ENTERED OTHER THAN BY CALL, indicating an attempt to return from a defined function which has not been called.

(*xvii*) INDIRECT REFERENCE THROUGH THE NULL STRING, indicating an attempt to use the null string as a name.

(*xviii*) OUT OF SPACE, indicating available storage has been exhausted.

(*xix*) SYSTEM ERROR, indicating a programming error in the SNOBOL3 interpreter, or a machine error.

4.2.2 *Post-Mortem Information*

On termination of program execution, information is printed for the programmer's use.

If execution was terminated as the result of an error, the number of the statement in which the error occurred and the current level of function call are printed in addition to the error message.

In either normal or error termination, statistics concerning execution are provided. The number of statements executed and the number of times the scanner was entered for pattern matching are tabulated. Storage allocation statistics are provided, and total millisecond times in the compiler and interpreter are given.

4.3 *Debugging Aids*

Several functions are available specifically for debugging.

4.3.1 *Function Tracing*

Function calls may be traced by use of TRACE(FLIST) where the value of FLIST is a list of function names for which a trace is desired. For example, the call

TRACE("SIZE,F,EQUALS")

results in the subsequent tracing of the three functions given. Both primitive and defined functions can be traced.

When a defined function being traced is called, a message is printed on the listing indicating the level from which the call was made, the name of the function and the value of all its arguments at the time of the call. When the function returns, a message is printed indicating the level to which the return is made, the name of the function, and the value returned. If a failure return is made, this is also indicated but no value is given.

When a primitive function being traced is called, a message is printed on the listing indicating the level at which the call was made, the name of the function, and the value of its arguments. If the function call returned successfully, the value is given. Otherwise failure is indicated.

The tracing of functions may be stopped by calling

STOPTR(FLIST)

where the value of FLIST is a list of functions for which tracing is to be stopped.

4.3.2 *String Tracing*

A name may be traced during execution by calling the function STRACE(NAME,FILE) which associates the value of NAME with the value of FILE in the trace sense. For example,

STRACE("Y","OUT")

would cause Y to be associated with logical file OUT. Subsequently every time a value is assigned to Y, a trace message will be printed on the associated file indicating the name being traced, its new value and the statement number where the value was assigned. STRACE is essentially an input-output association function and behaves like the other association functions. Consequently STRACE detaches NAME from any other input or output association. Similarly, if the second argument is null, association is made with the standard output file. String tracing may be terminated by detaching NAME.

4.3.3 *Diagnostics Resulting from Tracing*

Two nonfatal errors may occur as a result of tracing. Advisory diagnostic messages are printed for these cases.

(*i*) (name) HAS NOT BEEN DEFINED AND WILL NOT BE TRACED, indicating a request to trace an undefined function of the indicated name.

(*ii*) ** THE FOLLOWING TRACE OUTPUT HAS BEEN TRUNCATED, indicating that the printing of a string or function trace would exceed internal storage limitations, an assembly parameter set to about 600 characters. In this case the trace printout is truncated.

4.3.4 *String Dumps*

An alphabetical listing of all strings with nonnull values may be obtained on termination of execution. The MODE function is used to request this string dump.

MODE("DUMPERR") causes a string dump if execution is terminated by an error during execution. MODE("DUMP") causes a string dump following either normal or error termination. These calls of the MODE function must of course be made before execution is terminated.

V. ACKNOWLEDGMENT

The authors wish to express their appreciation to the many people who have contributed their ideas to the design of SNOBOL3. D. L. Clark, G. F. Faulhaber, G. K. Manacher, M. D. McIlroy, J. F. Poage, A. Simon, L. C. Varian, and D. Walters have been particularly helpful.

The assistance of G. K. Manacher and L. C. Varian in the implementation of the system is most gratefully acknowledged.

APPENDIX A

Catalog of Primitive Functions

Primitive functions may be divided into categories according to the nature of their operation and area of applicability. Individual functions are described in the appropriate sections. For reference purposes a complete list of primitive functions follow.

A. Numerical Functions (Section 3.2)
   1. .EQ(X,Y)
   2. .NE(X,Y)
   3. .LE(X,Y)
   4. .LT(X,Y)
   5. .GE(X,Y)
   6. .GT(X,Y)
   7. .NUM(X)
   8. .REMDR(X,Y)

B. Diagnostic Functions (Section 4.3)
   1. TRACE(FLIST)
   2. STOPTR(FLIST)
   3. STRACE(NAME,FILE)

C. Input-Output and File Manipulation Functions (Section 3.5)
   1. READ(NAME,FILE)
   2. PRINT(NAME,FILE)
   3. PUNCH(NAME,FILE)
   4. EJECT(NAME)
   5. REWIND(NAME)
   6. BSREC(NAME)
   7. DETACH(NAME)
   8. OPEN(KEY,FILE)

D. System Mode Functions
   1. MODE(X)
      a. "ANCHOR" — Sections 2.8.1 and 3.3.3
      b. "UNANCHOR" — Section 3.3.3
      c. "INTEGER" — Section 3.2.3
      d. "TRUNCATION" — Section 3.2.3
      e. "SYSTEM" — Section 3.4.4
      f. "ENDJOB" — Section 3.4.4
      g. "DUMP" — Section 4.3.4
      h. "DUMPERR" — Section 4.3.4
   2. ANCHOR( ) — Section 3.3.3
   3. UNANCH( ) — Section 3.3.3

E. Functions Relating to Functions
   1. DEFINE(FORM,LABEL,NAMES) — Sections 2.8.2 and 3.7

2. OPSYN(NEW,OLD)	Section 3.6.2
3. CALL(FNC)	Section 3.6.2
F. Miscellaneous Functions	
1. EQUALS(X,Y)	Sections 2.8.1 and 3.6.3
2. UNEQL(X,Y)	Section 3.6.3
3. TRIM(S)	Section 3.6.3
4. TIME( )	Section 3.6.3
5. DATE( )	Section 3.6.3
6. SIZE(S)	Sections 2.8.1 and 3.6.3
7. SYSTEM(FILE)	Section 3.4.4

APPENDIX B

This appendix contains three sample programs. These programs are designed to illustrate various uses and features of SNOBOL3.

The first two programs involve symbolic evaluations which are inherently recursive. The third is an example of text manipulation.

All three examples use pattern matching in various forms. Statement 20 in the first example illustrates the use of back referencing to determine whether two lists have an element in common. The second program illustrates function tracing.

```
*         EXAMPLE 1.  THE WANG ALGØRITHM
*
*              THIS PRØGRAM IS THE ALGØRITHM BY HAØ WANG (CF. "TØWARD
*         MECHANICAL MATHEMATICS", IBM JØURNAL ØF RESEARCH AND
*         DEVELØPMENT 4(1) JAN.1960 PP.2-22.)  FØR A PRØØF-DECISIØN
*         PRØCEDURE FØR THE PRØPØSITIØNAL CALCULUS.  IT PRINTS ØUT A
*         PRØØF ØR DISPRØØF ACCØRDING AS A GIVEN FØRMULA IS A THEØREM
*         ØR NØT.  THE ALGØRITHM USES SEQUENTS WHICH CØNSIST ØF TWØ
*         LISTS ØF FØRMULAS SEPARATED BY AN ARRØW (--*).  INITIALLY, FØR
*         A GIVEN FØRMULA F THE SEQUENT
*
*                   --* F
*
*         IS FØRMED.  WANG HAS DEFINED RULES FØR SIMPLIFYING A FØRMULA
*         IN A SEQUENT BY REMØVING THE MAIN CØNNECTIVE AND THEN
*         GENERATING A NEW SEQUENT ØR SEQUENTS.  THERE IS A TERMINAL
*         TEST FØR A SEQUENT CØNSISTING ØF ØNLY ATØMIC FØRMULAS:
*
*              A SEQUENT CØNSISTING ØF ØNLY ATØMIC FØRMULAS IS VALID IF
*              THE TWØ LISTS ØF FØRMULAS HAVE A FØRMULA IN CØMMØN.
*
*         BY REPEATED APPLICATIØN ØF THE RULES ØNE IS LED TØ A SET ØF
*         SEQUENTS CØNSISTING ØF ATØMIC FØRMULAS.  IF EACH ØNE ØF THESE
*         SEQUENTS IS VALID THEN SØ IS THE ØRIGINAL FØRMULA.



          DEFINE("WANG(ANTE,CØNSEQ)" , "WANG","PHI,PSI")
*         READ IN THE EXPRESSIØN

READ      EXP        =  TRIM(SYSPIT)                /F(END)
          SYSPØT     =
          SYSPØT     =  "FØRMULA: "  EXP
          SYSPØT     =
          WANG(," "  EXP)                           /F(INVALID)
          SYSPØT     =  "VALID"                     /(READ)
INVALID   SYSPØT     =  "NØT VALID"                 /(READ)
WANG      SYSPØT     =  ANTE   " --* "  CØNSEQ
          ANTE       " NØT("  *(PHI)*  ")" =        /S(ANØT)
          ANTE       " AND(" *(PHI)* "," *(PSI)* ")" =    /S(AAND)
          ANTE       " IMP(" *(PHI)* "," *(PSI)* ")" =    /S(AIMP)
          ANTE       " ØR(" *(PHI)* "," *(PSI)* ")" =     /S(AØR)
          ANTE       " EQU(" *(PHI)* "," *(PSI)* ")" =    /S(AEQU)

          CØNSEQ     " NØT(" *(PHI)* ")" =                /S(CNØT)
          CØNSEQ     " AND(" *(PHI)* "," *(PSI)* ")" =    /S(CAND)
          CØNSEQ     " IMP(" *(PHI)* "," *(PSI)* ")" =    /S(CIMP)
          CØNSEQ     " ØR(" *(PHI)* "," *(PSI)* ")" =     /S(CØR)
          CØNSEQ     " EQU(" *(PHI)* "," *(PSI)* ")" =    /S(CEQU)

          (ANTE " :" CØNSEQ " ")    " " *(P)* " " ** ":" ** " " P " "
                                                    /S(RETURN)F(FRETURN)

ANØT      WANG(ANTE,CØNSEQ  " "  PHI)               /S(RETURN)F(FRETURN)

AAND      WANG(ANTE  " "  PHI  " "  PSI, CØNSEQ)    /S(RETURN)F(FRETURN)

AØR       WANG(ANTE  " "  PHI,CØNSEQ)               /F(FRETURN)
          WANG(ANTE  " " PSI,CØNSEQ)                /S(RETURN)F(FRETURN)

AIMP      WANG(ANTE  " "  PSI,CØNSEQ)               /F(FRETURN)
          WANG(ANTE,CØNSEQ  " " PHI)                /S(RETURN)F(FRETURN)

AEQU      WANG(ANTE  " " PHI  " "  PSI,CØNSEQ)      /F(FRETURN)
          WANG(ANTE,CØNSEQ  " " PHI  " "  PSI)      /S(RETURN)F(FRETURN)

CNØT      WANG(ANTE  " "  PHI,CØNSEQ)               /S(RETURN)F(FRETURN)

CAND      WANG(ANTE,CØNSEQ  " " PHI)                /F(FRETURN)
          WANG(ANTE,CØNSEQ  " "  PSI)               /S(RETURN)F(FRETURN)

CØR       WANG(ANTE,CØNSEQ  " " PHI  " "  PSI)      /S(RETURN)F(FRETURN)

CIMP      WANG(ANTE  " "  PHI,CØNSEQ  " "  PSI)     /S(RETURN)F(FRETURN)

CEQU      WANG(ANTE  " "  PHI,CØNSEQ  " "  PSI)     /F(FRETURN)
          WANG(ANTE  " "  PSI,CØNSEQ  " "  PHI)     /S(RETURN)F(FRETURN)
END

SUCCESSFUL CØMPILATIØN
```

```
FØRMULA: IMP(NØT(ØR(P,Q)),NØT(P))

 --*  IMP(NØT(ØR(P,Q)),NØT(P))
 NØT(ØR(P,Q)) --*  NØT(P)
 --*  NØT(P) ØR(P,Q)
 P --*  ØR(P,Q)
 P --*  P Q
VALID

FØRMULA: IMP(AND(NØT(P),NØT(Q)),EQU(P,Q))

 --*  IMP(AND(NØT(P),NØT(Q)),EQU(P,Q))
 AND(NØT(P),NØT(Q)) --*  EQU(P,Q)
 NØT(P) NØT(Q) --*  EQU(P,Q)
 NØT(Q) --*  EQU(P,Q) P
 --*  EQU(P,Q) P Q
 P --*  P Q Q
 Q --*  P Q P
VALID

FØRMULA: IMP(IMP(ØR(P,Q),ØR(P,R)),ØR(P,IMP(Q,R)))

 --*  IMP(IMP(ØR(P,Q),ØR(P,R)),ØR(P,IMP(Q,R)))
 IMP(ØR(P,Q),ØR(P,R)) --*  ØR(P,IMP(Q,R))
 ØR(P,R) --*  ØR(P,IMP(Q,R))
 P --*  ØR(P,IMP(Q,R))
 P --*  P IMP(Q,R)
 P Q --*  P R
 R --*  ØR(P,IMP(Q,R))
 R --*  P IMP(Q,R)
 R Q --*  P R
 --*  ØR(P,IMP(Q,R)) ØR(P,Q)
 --*  ØR(P,Q) P IMP(Q,R)
 Q --*  ØR(P,Q) P R
 Q --*  P R P Q
VALID

.................................................................................................................

 NØRMAL EXIT FRØM SNØBØL AT LEVEL       0

SNØBØL RUN STATISTICS , NØ. ØF RULES EXECUTED =   254 NØ ØF SCANNER ENTRIES =   187

STØRAGE ALLØCATIØN STATISTICS  --     171  STRINGS STØRED     397  WØRDS FØR STØRED STRINGS
        600 REFERENCE ASSIGNMENT WØRDS ,      0 REFERENCE,      0   GARBAGE, AND
        0 HYPER-GARBAGE CØLLECTIØN(S)

ELAPSED TIMES-CØMPILER    620 , INTERPRETER       735 IN MS
```

```
*        EXAMPLE 2.  DIFFERENTIATIØN ØF ALGEBRAIC EXPRESSIØNS
*
*             THIS PRØGRAM DIFFERENTIATES A FULLY PARENTHESIZED
*        ALGEBRAIC EXPRESSIØN WITH RESPECT TØ X.  THE EXPØNENTIATIØN
*        ØPERATØR IS INDICATED BY A DØLLAR SIGN.

         DEFINE("D(D)","DO","U,V,N")                                          1
         DEFINE("SIMPLIFY(EXP)","SO","U")                                     2
         TRACE("D,SIMPLIFY")                                                  3

*        READ IN THE EXPRESSIØN.

READ     EXP         =  TRIM(SYSPIT)                    /F(END)               4

         MØDE("ANCHØR")                                                       5
         SYSPØT      =                                                        6
         SYSPØT      =  "THE DERIVATIVE ØF " EXP " IS " SIMPLIFY(D(EXP))      7
                                                        /(END)                7
.

*        THE FUNCTIØN "D"

DO       D           "(" *(U)* "+" *(V)* ")" = "(" D(U) "+" D(V) ")"          8
                                                        /S(RETURN)            8
.
         D           "(" *(U)* "-" *(V)* ")" = "(" D(U) "-" D(V) ")"          9
.                                                       /S(RETURN)            9

         D           "(" *(U)* "*" *(V)* ")" = "((" U "*" D(V) ")+("         10
.                              V  "*" D(U) "))"         /S(RETURN)           10

         D           "(" *(U)* "/" *(V)* ")" = "((((" V "*" D(U) ")-("       11
.                          U "*" D(V) "))/(" V "$2))"   /S(RETURN)           11

         D           "(" *(U)* "$" *(N)* ")" =                               12
.                              "((" N "*(" U "$" N - "1" "))*" D(U) ")"      12
.                                                       /S(RETURN)           12

         D           =  EQUALS(D,"X")  "1"              /S(RETURN)           13
         D           =  "0"                             /(RETURN)            14

*        THE FUNCTIØN "SIMPLIFY"

S0       MØDE("UNANCH")                                                      15
S1       EXP         "(" *(U)* "*0)"   =  "0"           /S(S1)               16
         EXP         "(0*" *(U)* ")"   =  "0"           /S(S1)               17
S2       EXP         "(" *(U)* "$1)"   =  U             /S(S2)               18
S3       EXP         "(" *(U)* "*1)"   =  U             /S(S3)               19
S4       EXP         "(1*" *(U)* ")"   =  U             /S(S4)               20
S5       EXP         "(0+" *(U)* ")"   =  U             /S(S5)               21
S6       EXP         "(" *(U)* "+0)"   =  U             /S(S6)               22
S7       EXP         "(" *(U)* "-0)"   =  U             /S(S7)               23
         SIMPLIFY    =  EXP                             /(RETURN)            24
END                                                                          25

SUCCESSFUL CØMPILATIØN
```

```
0 LEVEL CALL ØF D("(((A*(X$2))+(B*X))+C)")

1 LEVEL CALL ØF D("((A*(X$2))+(B*X))")

2 LEVEL CALL ØF D("(A*(X$2))")

3 LEVEL CALL ØF D("(X$2)")

4 LEVEL CALL ØF D("X")

4 LEVEL RETURN ØF D = "1"

3 LEVEL RETURN ØF D = "((2*(X$1))*1)"
```

```
3 LEVEL CALL ØF D("A")

3 LEVEL RETURN ØF D = "0"

2 LEVEL RETURN ØF D = "((A*((2*(X$1))*1))+((X$2)*0))"

2 LEVEL CALL ØF D("(B*X)")

3 LEVEL CALL ØF D("X")

3 LEVEL RETURN ØF D = "1"

3 LEVEL CALL ØF D("B")

3 LEVEL RETURN ØF D = "0"

2 LEVEL RETURN ØF D = "((B*1)+(X*0))"

1 LEVEL RETURN ØF D = "(((A*((2*(X$1))*1))+((X$2)*0))+((B*1)+(X*0)))"

1 LEVEL CALL ØF D("C")

1 LEVEL RETURN ØF D = "0"

0 LEVEL RETURN ØF D = "((((A*((2*(X$1))*1))+((X$2)*0))+((B*1)+(X*0)))+0)"

0 LEVEL CALL ØF SIMPLIFY("((((A*((2*(X$1))*1))+((X$2)*0))+((B*1)+(X*0)))+0)")

0 LEVEL RETURN ØF SIMPLIFY = "((A*(2*X))+B)"
 THE DERIVATIVE ØF (((A*(X$2))+(B*X))+C) IS ((A*(2*X))+B)

------------------------------------------------------------------------------------------------------------------------

 NØRMAL EXIT FRØM SNØBØL AT LEVEL        0

SNØBØL RUN STATISTICS , NØ. ØF RULES EXECUTED =     71 NØ ØF SCANNER ENTRIES =     54

STØRAGE ALLØCATIØN STATISTICS  --      125  STRINGS STØRED      216  WØRDS FØR STØRED STRINGS
        450 REFERENCE ASSIGNMENT WØRDS ,       0 REFERENCE,       0   GARBAGE, AND
        0 HYPER-GARBAGE CØLLECTIØN(S)

ELAPSED TIMES-CØMPILER     474 , INTERPRETER       380 IN MS
```

```
*         EXAMPLE 3.  EDITØR
*
*              THIS PRØGRAM READS IN TEXT AND PRINTS N CHARACTERS
*         PER LINE.  EXTRA BLANKS ARE INSERTED BETWEEN WØRDS
*         TØ FILL ØUT LINES.

          MØDE("ANCHØR")                                                        1
          DEFINE("INSERT(K,LINE)","IN","BLANK,WØRD")                            2

*         READ IN N
          TRIM(SYSPIT)         *N*                                              3

READ      TEXT   =  TEXT TRIM(SYSPIT) " "            /S(READ)                   4
          SYSPØT    =                                                           5
          SYSPØT    =  "PRINT THE FØLLØWING STRING " N                          6
.                         " CHARACTERS PER LINE:"                               6
          SYSPØT    =                                                           7

*         PRINT THE INPUT TEXT

          TEXT1   =   TEXT                                                      8
PRINT     TEXT1   *SYSPØT/"90"*    =                 /S(PRINT)                  9
          SYSPØT    =   TEXT1                                                  10
          SYSPØT    =                                                          11
          SYSPØT    =                                                          12
TEST      .GT(SIZE(TEXT),N)                          /F(LASTLINE)              13
          K         =      "0"                                                 14
SCAN      TEXT      *LINE/(N - K)*  " "  =           /F(BUMPK)                 15
          SYSPØT    =  INSERT(K,LINE)                /(TEST)                   16
BUMPK     K         =    .LT(K,N)  K + "1"           /S(SCAN)F(ERR)            17

*    FUNCTIØN TØ INTERSPERSE K BLANKS IN A LINE

IN        INSERT    =  .EQ(K,"0")  LINE              /S(RETURN)                18
          LINE      **  " "                          /S(BLINK)                 19
          INSERT    =  LINE                          /(RETURN)                 20
BLINK     BLANK     =  BLANK  " "                                              21
LØØP      LINE     *WØRD*  BLANK   =                 /F(MØRE)                  22
          INSERT    =   INSERT  WØRD BLANK  " "                                23
          K         =    .GT(K,"1")    K - "1"       /S(LØØP)                  24
          INSERT    =  INSERT LINE                   /(RETURN)                 25
MØRE      LINE      =  INSERT LINE                                             26
          INSERT    =                                /(BLINK)                  27

LASTLINE  SYSPØT    =  TEXT                          /(END)                    28
ERR       SYSPØT    =                                                          29
          SYSPØT    = "** EDITØR CANNØT PRINT LINE BECAUSE N IS TØØ SMALL."    30
END                                                                            31

SUCCESSFUL CØMPILATIØN

-----------------------------------------------------------------------------------

PRINT THE FØLLØWING STRING 75 CHARACTERS PER LINE:

TWØ FUNCTIØNS HAVE BEEN ADDED TØ INCREASE THE FLEXIBILITY ØF DEALING WITH THE SYSTEM INPUT
 SØURCE. GETSRC RETURNS AS VALUE THE CURRENT SYSTEM SØURCE. SETSRC SETS THE CURRENT SYSTEM
 SØURCE TØ THE VALUE ØF FILE. A NULL VALUE IS RETURNED. THREE I/Ø FUNCTIØNS HAVE BEEN ADDE
D WHICH TAKE FILES AS ARGUMENTS. THESE FUNCTIØNS CØMPLEMENT THE CØRRESPØNDING SNØBØL3 FUNC
TIØNS WHICH REQUIRE NAMES ASSØCIATED WITH FILES AS ARGUMENTS.

TWØ  FUNCTIØNS  HAVE BEEN ADDED TØ INCREASE THE FLEXIBILITY ØF DEALING WITH
THE SYSTEM INPUT SØURCE. GETSRC RETURNS AS VALUE THE CURRENT SYSTEM SØURCE.
SETSRC SETS THE CURRENT SYSTEM SØURCE TØ THE VALUE ØF FILE. A NULL VALUE IS
RETURNED.  THREE  I/Ø  FUNCTIØNS  HAVE  BEEN  ADDED  WHICH  TAKE  FILES  AS
ARGUMENTS.  THESE  FUNCTIØNS CØMPLEMENT THE CØRRESPØNDING SNØBØL3 FUNCTIØNS
WHICH REQUIRE NAMES ASSØCIATED WITH FILES AS ARGUMENTS.

-----------------------------------------------------------------------------------

 NØRMAL EXIT FRØM SNØBØL AT LEVEL        0

SNØBØL RUN STATISTICS , NØ. ØF RULES EXECUTED =    130 NØ ØF SCANNER ENTRIES =     42

STØRAGE ALLØCATIØN STATISTICS --      194  STRINGS STØRED     1296  WØRDS FØR STØRED STRINGS
        600 REFERENCE ASSIGNMENT WØRDS ,       0 REFERENCE,     0   GARBAGE, AND
        0 HYPER-GARBAGE CØLLECTIØN(S)

ELAPSED TIMES-CØMPILER    410 , INTERPRETER       547 IN MS
```

REFERENCES

1. Newell, A., Ed., *Information Processing Language-V Manual,* Prentice-Hall, 1961.
2. McCarthy, J., Recursive Functions of Symbolic Expressions and Their Computation by Machine, Part I. Comm., *ACM 3* April, 1960, p. 184.
3. An Introduction to COMIT Programming, The Research Lab of Electronics and the Computation Center, M.I.T., 1961.
4. Farber, D. J., Griswold, R. E., and Polonsky, I. P., SNOBOL, a String Manipulation Language, J. ACM, *11,* No. 1, 1964.
5. Manacher, G. K., A Package of Subroutines for the SNOBOL Language. (Unpublished)
6. Griswold, R. E. and Polonsky, I. P., Tree Functions for SNOBOL3. (Unpublished)
7. Griswold, R. E., Linked-List Functions for SNOBOL3. (Unpublished)
8. Farber, D. J., Griswold, R. E., Manacher, G. K., Polonsky, I. P., and Varian, L. C., Programming Machine-Language Functions for SNOBOL3. (Unpublished)

Acoustic Scattering of Light in a Fabry-Perot Resonator

By M. G. COHEN and E. I. GORDON

(Manuscript received March 25, 1966)

Light scattering in a Fabry-Perot cavity by an acoustic beam is discussed. An heuristic treatment based on momentum conservation is used to determine the conditions, acoustic bandwidth and enhancement of the scattering interaction. A more detailed and rigorous calculation based on a coupled-mode formalism is also described. Experimental results using fused quartz Fabry-Perot cavities, single-frequency 6328Å light and acoustic waves in the frequency range 200–500 Mc/s, are presented. Enhancement relative to single-pass scattering by a factor of 50 is easily achieved. Modulation depths of twenty-five percent with a bandwidth of several megacycles have been observed.

I. INTRODUCTION

When a collimated light beam traverses a collimated high-frequency acoustic beam, it is possible for the acoustic beam to scatter light into a single, well-defined angle (into a number of well-defined angles when the acoustic frequency is sufficiently low). The amount of light scattered depends critically on the angle of the light beam relative to the acoustic wavefront (the Bragg angle) and is roughly proportional to the square of the acoustic beam width for constant acoustic intensity.[1] While it is possible to scatter all of the incident light, this usually requires impractically large amounts of acoustic power, especially at frequencies greater than a few tens of megacycles/second.

Acoustic scattering of light has considerable experimental interest since it allows a convenient method of probing transparent media to measure such things as elastic and photoelastic constants and acoustic loss or phonon lifetimes. These measurements can be made using thermally generated (Brillouin scattering)[2,3] or externally generated sound.[4] In the latter case, the scattering also allows determination of the acoustic beam shape and direction. In addition, the scattering interaction gives rise to a class of light modulation and deflection devices.[5,6] For amplitude modulation purposes, the bandwidth of acoustic devices corresponds to the transit time of the sound across the waist of a light beam whose diffraction angle equals the diffraction angle of the sound.[5] Interest in these devices and similar ones using slow electromagnetic waves in electro-optic materials[6] would be more than academic if a sizeable percentage of the incident light could be conveniently deflected.

Resonating the acoustic medium could enhance the strength of the interaction but would excessively limit the bandwidth or transient response. On the other hand, the large velocity of light allows resonating the interaction region optically with less serious consequences to the bandwidth. Such a technique would be an extension of what has already been described for electro-optic modulation. In the Fabry-Perot electro-optic modulator,[7] the electro-optic standing wave serves to couple energy from one axial Fabry-Perot mode excited by the incident light to other axial modes. In the acoustic case, the coupling is necessarily between off-axis modes of the resonator.

Figs. 1 and 2 illustrate the arrangement required to achieve a resonant scattering interaction. For the sake of discussion, it is assumed that the incident light beam is very wide so that the angle of incidence is well defined. Under ideal conditions the sound of angular frequency Ω and velocity v travels in a plane parallel to the mirrors; unavoidably the sound travels at a very small angle Ψ relative to the mirrors as shown in Fig. 2 and this case must be considered also. The incident light of angular frequency ω and velocity c' in the medium is set at an angle θ corresponding to an off-axis resonance of the cavity. This requires that

$$k \cos \theta = a\pi/L \tag{1}$$

in which $k = \omega/c'$, L is the mirror spacing and a is an integer. Now consider Fig. 1; the scattered light with propagation constant $k' = k(1 + \Omega/\omega)^*$ is sent into an angle θ' defined precisely by

$$k' \sin \theta' = k \sin \theta + K \tag{2}$$

in which $K = \Omega/v$ is the acoustic propagation constant; (2) follows from the requirement of momentum conservation. The various scattered beams arising from each pass of the light will add in-phase when the angle θ' corresponds to a resonance of the cavity at the frequency $\omega + \Omega$. Hence, the angle θ' must satisfy the requirement

$$k' \cos \theta' = a'\pi/L. \tag{3}$$

* For the geometry shown, the scattered light is upshifted in frequency to a value $\omega + \Omega$. It is assumed that optical dispersion is negligible over the frequency range of interest so that $k' = (\omega + \Omega)/c'$.

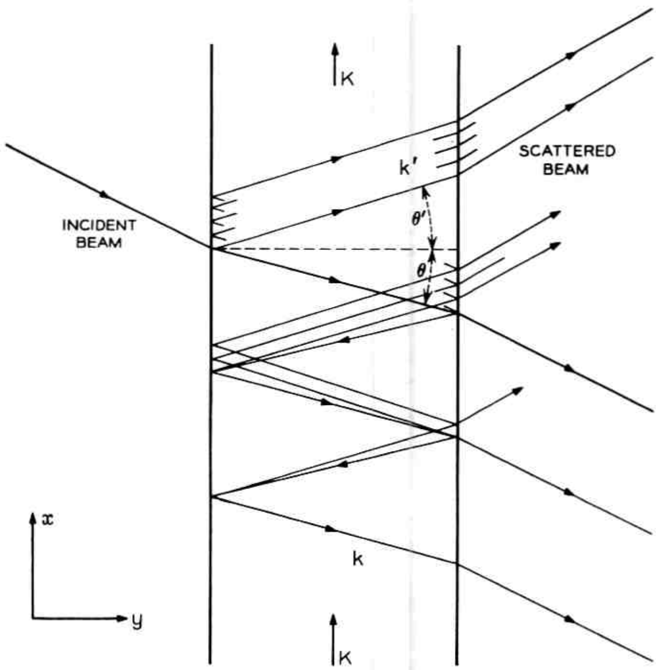

Fig. 1 — Schematic arrangement for acoustic scattering of light in an optically resonant geometry. In this case the acoustic wave travels parallel to the mirror planes.

When the acoustic beam has no y-directed momentum, momentum conservation in the y-direction (the Bragg condition) requires that

$$k' \cos \theta' = k \cos \theta \tag{4}$$

or

$$a' = a.$$

Since the acoustic beam is not extremely wide in practice, the transverse (y-directed) momentum or propagation constant is not precisely defined. This allows some deviation from the strict requirement

$$k' \cos \theta' = k \cos \theta,$$

and it is not absolutely necessary that $a' = a$. However, (4) represents the condition for optimum scattering.

Following the multiply-reflected beam, it can be seen that the Bragg condition is satisfied along each leg of the round trip, hence energy is efficiently scattered into the mode at angle θ' on each pass through the

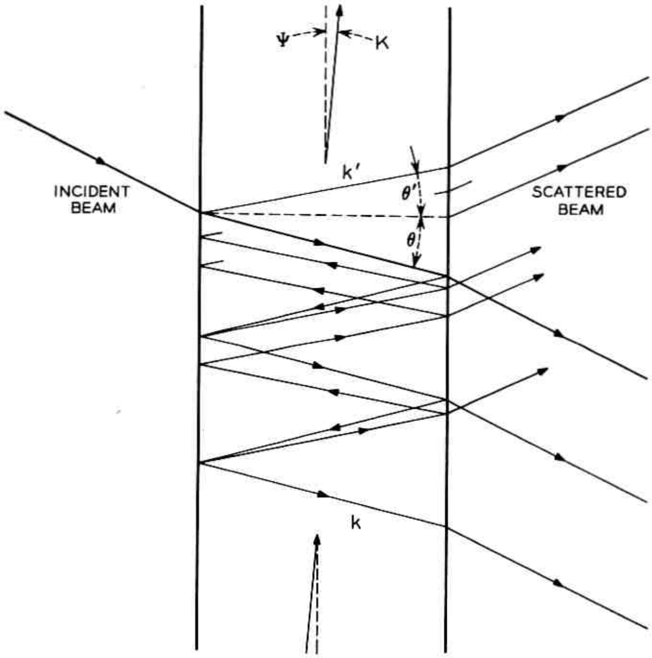

Fig. 2 — Schematic arrangement for acoustic scattering of light in an optically resonant geometry. In this case the acoustic wave travels at a small angle Ψ relative to the mirror planes.

acoustic beam. Since the effective number of passes is given approximately by $R^{\frac{1}{2}}(1 - R)^{-1}$, in which R is the mirror reflectivity,[7] the scattered optical energy is nominally increased over the single-pass value by the factor $R(1 - R)^{-2}$; for $R = 0.9$, $R(1 - R)^{-2} = 90$. The presence of loss or a finite aperture in the Fabry-Perot cavity will, of course, reduce the enhancement or gain in scattered energy.

When the acoustic beam moves at an angle Ψ as shown in Fig. 2, the situation is somewhat more complicated. The acoustic frequency must be chosen so that the angle of the scattered radiation defined by

$$k' \sin \theta' = k \sin \theta + K \cos \Psi \tag{5}$$

falls into a Fabry-Perot mode and the Bragg condition requires that

$$\begin{aligned} \pm K \sin \Psi &\approx k \cos \theta - k' \cos \theta' \\ &= (a - a')\pi/L. \end{aligned} \tag{6}$$

The requirement for $\Psi \neq 0$ can be well satisfied for either the plus or the minus sign and hence for only one leg of the round trip as shown in Fig. 2. Thus, the enhancement of the path length is reduced by $\frac{1}{2}$ and the scattered power by $\frac{1}{4}$ over that represented by the case $\Psi = 0$, (Fig. 1).

When the acoustic frequency Ω is chosen according to (5) the energy scattered from the mode excited by the incident beam will fall precisely into the optimum angle θ' for the scattered mode. When the acoustic frequency deviates from the proper value the scattering angle will vary and the scattered intensity will decrease. With the advantage of hindsight, one expects that the amount of the decrease will be determined by the angular halfwidth of the Fabry-Perot mode. For a lossless Fabry-Perot cavity the angular dependence of the transmitted intensity T is given by the Haidinger fringe formula[8] and can be written

$$T(\Omega) = \frac{1}{1 + [4R/(1 - R)^2] \sin^2 k'L\,[\cos\theta'(\Omega) - \cos\theta'(\Omega_o)]} \tag{7}$$

in which $\theta'(\Omega)$ and $\theta'(\Omega_o)$ are defined by (5) for different values of K (defined as Ω/v and Ω_o/v; Ω_o being the optimum acoustic frequency). Using (5), $T(\Omega)$ can be written finally as

$$T(\Omega) \approx \frac{1}{1 + [4R/(1 - R)^2] \sin^2 (\Omega - \Omega_o)\,(L\cos\Psi/v)\tan\theta'(\Omega_o)} \tag{8}$$

which predicts an acoustic bandwidth

$$\Delta\Omega/2\pi \approx v(1 - R)R^{-\frac{1}{2}}/2\pi L\cos\Psi\,|\tan\theta'(\Omega_o)|\,. \tag{9}$$

More detailed calculations indicate that in the limit of no loss (9) is precisely correct. An additional factor $1 + |\tan\theta'/\tan\theta|$ will appear when the angular spread of the light in the mode excited by the incident light beam of finite width is taken into account.

It may be shown by forming the product of the fraction of the incident light scattered or scattering efficiency and the bandwidth that the resonant Fabry-Perot device has a figure of merit identical to that of the nonresonant or single-pass modulation device.[9] The Fabry-Perot device has practical interest only when efficient narrow-band (<2 megacycle/second) devices are required since nonresonant modulators are difficult to optimize for narrow bandwidths.

The preceding discussion indicates in a qualitative way the properties of a scattering interaction in a Fabry-Perot cavity. In the following section, a more rigorous theory of the scattering interaction will be given which validates the simple model described above. This is followed by a description of experiments giving results in substantial agreement with the calculated results.

II. THEORY

The model chosen for discussion is that of an open Fabry-Perot cavity with unlimited x-dimension as shown in Fig. 3. The light beam is incident at an angle θ_a, at a frequency ω_a corresponding to a cavity resonance defined by

$$k_a \cos \theta_a = a\pi/L \tag{10}$$

in which $k_a = \omega_a/c'$. The cavity electric field has the form

$$\mathbf{E}(x,y,t) = \hat{k}[\exp + \alpha_a x][\exp i(\omega_a t + k_a \sin \theta_a x)] \sin (a\pi y/L) \tag{11}$$

corresponding to a decaying wave traveling to $x = -\infty$. The field is resonant in the y-dimension. The exponential loss, with loss parameter α_a, arises from the mirror reflection loss as well as internal losses in the resonator. It is shown in Appendix A that α_a is given by

$$\alpha_a = \tfrac{1}{2}k_a/Q_{L,a} \mid \sin \theta_a \mid \tag{12}$$

in which $Q_{L,a}$ is the loaded Q of the cavity (defined later). The E field is assumed to be polarized in the z-direction (denoted by $\hat{k}$).

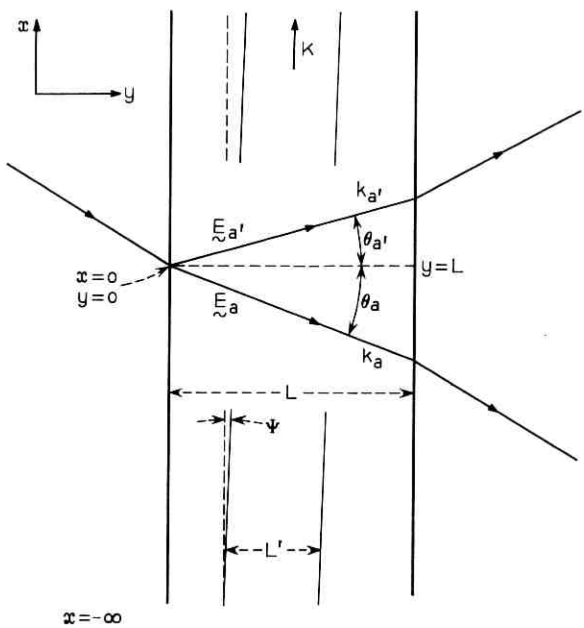

Fig. 3 — Scattering interaction geometry for coupled-mode calculation.

The eigenfunctions for the cavity, normalized in the half space $-\infty \leqq x \leqq 0$, can be written

$$\mathbf{E}_a(\omega_a) = \hat{k}(4\alpha_a/L)^{\frac{1}{2}} \sin (a\pi y/L) \exp (\alpha_a + ik_a \sin \theta_a)x. \quad (13)$$

The coupled-mode equations for the Fabry-Perot resonator are given in Ref. 7, and can be written

$$\begin{aligned} &\partial^2 e_a/\partial t^2 + \omega_a^2 e_a + (\omega_a/Q_{L,a})\partial e_a/\partial t \\ &\qquad + \partial^2[\langle a \mid \delta\varepsilon/\varepsilon \mid a' \rangle^* e_{a'}]/\partial t^2 \\ &\qquad\qquad = -\tfrac{1}{2}(\omega_a/Q_a)\partial e(\omega_a t)/\partial t \\ &\partial^2 e_{a'}/\partial t + \omega_{a'}^2 e_{a'} + (\omega_{a'}/Q_{L,a'})\partial e_{a'}/\partial t \\ &\qquad + \partial^2[\langle a \mid \delta\varepsilon/\varepsilon \mid a' \rangle e_a]/\partial t^2 \\ &\qquad\qquad = 0. \end{aligned} \quad (14)$$

Here $e_a(t)$ and $e_{a'}(t)$ are the amplitudes of the incident-beam and scattered-beam modes. The total field is given by

$$\mathbf{E} = e_a\mathbf{E}_a + e_{a'}\mathbf{E}_{a'} .$$

Note that only two modes are coupled here; any energy scattered from mode a' into an angle different from θ_a cannot correspond to a Fabry-Perot mode. (In general, the angles of the Fabry-Perot rings increase as the square root of an integer while the acoustic scattering angles increase as an integer and momentum conservation can occur for only one angle.) The quantity Q_a is the coupling Q defined by[7]

$$\begin{aligned} Q_a &= \tfrac{1}{4}\pi a(1 + R^{\frac{1}{2}})^2/(1 - R) \cos^2 \theta_a \\ &\approx \pi a R^{\frac{1}{2}}/(1 - R) \cos^2 \theta_a \dagger \end{aligned} \quad (15)$$

and $Q_{L,a}$ is the loaded Q defined by[7]

$$Q_{L,a}^{-1} = Q_a^{-1} + Q_d^{-1} \quad (16)$$

in which Q_d is the dielectric Q of the optical medium. The coupling of the Fabry-Perot modes is defined by the integral

$$\langle a \mid \delta\varepsilon/\varepsilon \mid a' \rangle = \int_{-\infty}^{0} dx \int_0^L dy\, \mathbf{E}_a \cdot [\delta\varepsilon(x,y,t)/\varepsilon]\mathbf{E}_{a'}^* \quad (17)$$

† The definition of Q_a given here differs from that of Ref. 7 in the appearance of the term $\cos^2 \theta_a$ which has the value unity for the modes considered there. The reason for this term can be understood simply by noting that $\pi a/\cos^2 \theta_a = k_a L/\cos \theta_a$ and $L/\cos \theta_a$ is the increased length over which energy is stored for the same reflection loss. Thus, the Q must be enhanced by this factor. It is also assumed that $1 + R^{\frac{1}{2}} \approx 2R^{\frac{1}{2}}$ which neglects terms of order $(1 - R)^2 \ll 1$.

in which ε is the unperturbed dielectric constant of the medium and $\delta\varepsilon$ is the perturbation produced by the sound. This can be written

$$\delta\varepsilon(x,y,t) = \delta\varepsilon \cos(\Omega t - Kx \cos\Psi - Ky \sin\Psi) \tag{18}$$

and can be written as a scalar quantity because of the choice of polarization axis. The acoustic beam travels at an angle Ψ relative to the mirrors, has a width L' and does not necessarily fill the space between mirrors. This introduces a filling factor L'/L into the expression for the coupling factor. The limits on the integration with respect to y in (17) will be taken as

$$\tfrac{1}{2}(L - L') \leqq y \leqq \tfrac{1}{2}(L + L').$$

Performing the integration, using (13) and (18) yields

$$\langle a \mid \delta\varepsilon/\varepsilon \mid a' \rangle = \tfrac{1}{2}\chi_{a,a'}(\Omega) \exp i\Omega t \tag{19}$$

with the coupling parameter $\chi_{a,a'}$ defined by

$$\chi_{a,a'}(\Omega) = \tfrac{1}{2}(\delta\varepsilon/\varepsilon)(L'/L)(i^{a-a'} \exp -i\tfrac{1}{2}(\Omega/v)L' \sin\Psi)$$
$$\times \left[\frac{\sin Y_+}{Y_+} + \frac{\sin Y_-}{Y_-}\right]\left[\frac{2\alpha_a^{\frac{1}{2}}\alpha_{a'}^{\frac{1}{2}}}{\alpha_a + \alpha_{a'}}\right] \Big/ [1 + iX]. \tag{20}$$

The parameter $Y_\pm$ is given by

$$Y_\pm = \tfrac{1}{2}[(a - a')\pi \pm (\Omega/v)L \sin\Psi]L'/L. \tag{21}$$

It can be shown that $\sin Y/Y$ is precisely the frequency dependence found for single-pass scattering when the angle of the incident light is held fixed. Hence, the requirement of small Y is nothing but the condition that the angle of incidence be close to the Bragg angle.[4] Note that when $a \neq a'$ and $(\Omega/v)L \sin\Psi \approx (a - a')\pi$ so that Y_+ is small and $\sin Y_+/Y_+ \approx 1$, then $Y_- \approx (a - a')\pi$ and $\sin Y_-/Y_- \approx 0$. This corresponds to the situation described earlier (Fig. 2) in which only one leg of the round trip scatters efficiently. However, when $a = a'$ then $Y_+ = -Y_- = KL' \sin\Psi$ which can be very small only when $\Psi \to 0$. In this case, $\sin Y_+/Y_+ + \sin Y_-/Y_- = 2$ corresponding to the condition of Fig. 1 wherein both legs of the round trip scatter efficiently.

Continuing the discussion of (20) the parameter X is given by

$$X = (k_a \mid \sin\theta_a \mid + k_{a'} \mid \sin\theta_{a'} \mid - K \cos\Psi)/(\alpha_a + \alpha_{a'}) \tag{22}$$

and is zero for some acoustic frequency $\Omega_{a,a'}$ defined implicitly by

$$k_a \mid \sin\theta_a \mid + k_a(1 + \Omega_{a,a'}/\omega_a) \mid \sin\theta_{a'} \mid - (\Omega_{a,a'}/v) \cos\Psi = 0 \tag{23}$$

subject to the constraint

$$k_a \cos\theta_a = a\pi/L$$
$$k_a(1 + \Omega_{a,a'}/\omega_a)\cos\theta_{a'} = a'\pi/L.$$

It follows that

$$\Omega_{a,a'} = \frac{[k_a v/\cos\psi][|\sin\theta_a| + |\sin\theta_{a'}|]}{[1 - (v/c')\sin\theta_{a'}/\cos\Psi]} \tag{24}$$

in which θ_a and $\theta_{a'}$ are the appropriate ring angles of the Fabry-Perot cavity.† The denominator in (24) is properly set equal to unity since $(v/c')\sin\theta_{a'} \sim \Omega_{a,a'}/\omega_a \ll 1$. Thus, $\Omega_{a,a'}$ represents the set of acoustic frequencies for which the scattered light will fall into a Fabry-Perot mode. There is one value of $\Omega_{a,a'}$ only for each pair of values a and a'. The parameter X may now be rewritten

$$\begin{aligned} X &= 2(\Omega_{a,a'} - \Omega)[1 + (v/c')\sin\theta_{a'}/\cos\Psi]\tau_{a,a'} \\ &\approx 2(\Omega_{a,a'} - \Omega)\tau_{a,a'} \end{aligned} \tag{25}$$

in which

$$\tau_{a,a'} = \tfrac{1}{2}(\cos\Psi)/v(\alpha_a + \alpha_{a'}) \tag{26}$$

corresponds to a transit time of the sound across an equivalent distance $\frac{1}{2}(\alpha_a + \alpha_{a'})^{-1}$, which is related to the decay distance of the light intensity in the cavity modes. The second term in the square bracket of (25) is neglected as discussed above. Using (10), (12), and (15), $\tau_{a,a'}$ can be written finally as

$$\tau_{a,a'} = \left(\frac{\cos\Psi}{v}\right)\left(\frac{L\tan|\theta_a|}{(1-R)R^{-\frac{1}{2}}}\right)\left[\frac{Q_a}{Q_{L,a}} + \frac{Q_{a'}}{Q_{L,a'}}\frac{\tan|\theta_a|}{\tan|\theta_{a'}|}\right]^{-1}. \tag{27}$$

The scattered amplitude can now be found by solving (14). Writing the mode amplitudes as

$$e_a = \hat{e}_a \exp i\omega_a t$$
$$e_{a'} = \hat{e}_{a'} \exp i(\omega_a + \Omega)t$$

and substituting into (14) yields

$$\begin{aligned} \hat{e}_{a'} &= \frac{-\frac{1}{2}i\chi_{a,a'}Q_{L,a'}\hat{e}_a}{1 - i2(\Omega - \Omega_{a,a'})Q_{L,a'}/\omega_a} \\ \hat{e}_a &= \frac{-\frac{1}{2}(Q_{L,a}/Q_a)e}{1 + \frac{1}{4}|\chi_{a,a'}|^2 Q_{L,a}Q_{L,a'}/(1 - i2(\Omega - \Omega_{a,a'})Q_{L,a'}/\omega_a)}. \end{aligned} \tag{28}$$

† $|\theta_{a'}|_{\omega_a \pm \Omega_{a,a'}}$ corresponding to a ring angle for an optical frequency $\omega_a \pm \Omega_{a,a'}$ can be related to the ring angle for frequency ω_a by the formula

$$|\sin\theta_{a'}|_{\omega_a \pm \Omega_{a,a'}} \approx |\sin\theta_{a'}|_{\omega_a}[1 \pm (\Omega_{a,a'}/\omega_a)\cot^2\theta_{a'}|_{\omega_a}].$$

Since $2Q_{L,a}/\omega_a$ is very small compared to $\tau_{a,a'}$ (the ratio is of order $(v/c')/\sin\theta_a$), the variation of $\hat{e}_a$ and $\hat{e}_{a'}$ with frequency arises almost entirely from the variation of $\chi_{a,a'}$ with frequency. Thus, the term $2(\Omega - \Omega_{a,a'})Q_{L,a'}/\omega_a$ will always be small compared to 1 in the frequency range of interest. The relative transmitted intensities in the incident and scattered modes are given by

$$\begin{aligned} T_a/T_{ao} &= [1 + \tfrac{1}{4} \mid \chi_{a,a'} \mid^2 Q_{L,a}Q_{L,a'}]^{-2} \\ T_{a'}/T_{ao} &= \tfrac{1}{4} \mid \chi_{a,a'} \mid^2 Q_{L,a'}{}^2/[1 + \tfrac{1}{4} \mid \chi_{a,a'} \mid^2 Q_{L,a}Q_{L,a'}]^2 \end{aligned} \tag{29}$$

in which T_{ao} is the transmission factor of the ath mode for $\chi_{a,a'} = 0$,

$$T_{ao} = (Q_{L,a}/Q_a)^2. \tag{30}$$

The relative scattered intensity in the absence of mirrors (single-pass) is given by[4]

$$\tfrac{1}{4}(\delta\varepsilon/\varepsilon)^2(k_aL')^2/\cos^2\theta_a\,, \tag{31}$$

while with the mirrors it is given by (for small $\chi_{a,a'}$)

$$\tfrac{1}{4} \mid \chi_{a,a'} \mid^2 Q_{L,a'}{}^2 T_{ao}\,. \tag{32}$$

Under optimum conditions,

$$\mid \chi_{a,a'} \mid = (\delta\varepsilon/\varepsilon)(L'/L)2\alpha_a^{\frac{1}{2}}\alpha_{a'}^{\frac{1}{2}}/(\alpha_a + \alpha_{a'})$$

so that the optimum gain or enhancement in scattered power over that obtained in single-pass scattering can be written using (12),

$$\begin{aligned} g_{\text{opt}} = 4\,\frac{T_{a'o}T_{ao}}{R^{-1}(1-R)^2}\Bigg[\left(\frac{Q_{L,a}}{Q_{L,a'}}\,\frac{\sin\mid\theta_a\mid}{\sin\mid\theta_{a'}\mid}\right)^{\frac{1}{2}} \\ + \left(\frac{Q_{L,a'}}{Q_{L,a}}\,\frac{\sin\mid\theta_{a'}\mid}{\sin\mid\theta_a\mid}\right)^{\frac{1}{2}}\Bigg]^{-2}. \end{aligned} \tag{33}$$

Note that for the case $a = a'$, neglecting transmission loss, $g_{\text{opt}} = R/(1-R)^2$ as expected. Under less than optimum conditions

$$\begin{aligned} g(\Omega) = g_{\text{opt}}\,\frac{1}{4}\Bigg[\frac{\sin\frac{1}{2}[(a-a')\pi + (\Omega/v)L\sin\Psi]L'/L}{\frac{1}{2}[(a-a')\pi + (\Omega/v)L\sin\Psi]L'/L} \\ + \frac{\sin\frac{1}{2}[(a-a')\pi - (\Omega/v)L\sin\Psi]L'/L}{\frac{1}{2}[(a-a')\pi - (\Omega/v)L\sin\Psi]L'/L}\Bigg]^2 \\ \times\frac{1}{1+[2(\Omega-\Omega_{a,a'})\tau_{a,a'}]^2} \end{aligned} \tag{34}$$

with $\tau_{a,a'}$ given by (27). A typical curve of $g(\Omega)$ is given in Fig. 4. For

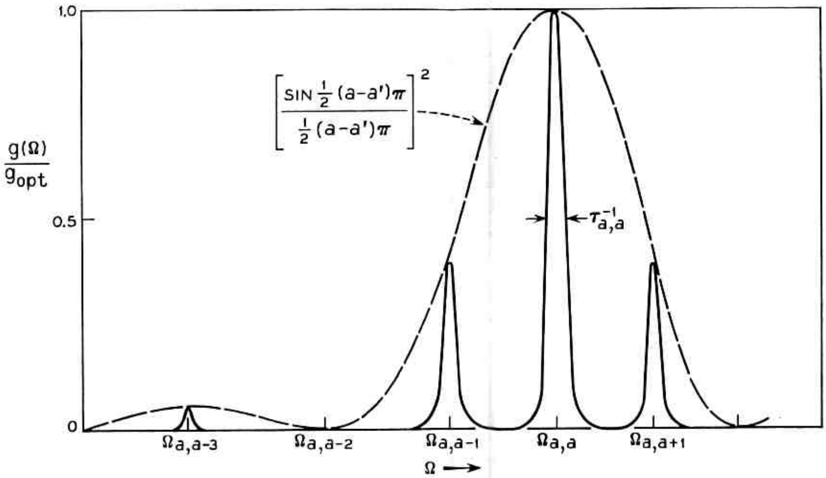

Fig. 4 — A plot of $g(\Omega)$, the enhancement factor, as a function of frequency, for the condition $\Psi = 0$ and $L = L'$.

this figure it has been assumed that $\Psi = 0$ and $L = L'$, so that

$$g(\Omega) = g_{\text{opt}} \frac{[\sin \frac{1}{2}(a - a')\pi/\frac{1}{2}(a - a')\pi]^2}{1 + [2(\Omega - \Omega_{a,a'})\tau_{a,a'}]^2}. \tag{35}$$

The maximum enhancement occurs for $a = a'$. The quantity $\tau_{a,a'}$ defined in (27) can be rewritten in terms of T_{ao} as

$$\tau_{a,a'} = \frac{\left(\dfrac{\cos\Psi}{v}\right)\left(\dfrac{L \tan|\theta_a|}{(1-R)R^{-\frac{1}{2}}}\right) T_{ao}^{\frac{1}{2}}}{1 + (T_{ao}/T_{a'o})^{\frac{1}{2}} \,|\tan\theta_a/\tan\theta_{a'}|}. \tag{36}$$

The acoustic bandwidth is given by

$$\Delta\Omega/2\pi = (2\pi\tau_{a,a'})^{-1}. \tag{37}$$

The similarity to (9) is apparent.

III. EXPERIMENT

In Section II expressions were derived for the optimum frequencies for acoustic scattering, enhancement factors, and the frequency width of the scattering interaction. Experiments were performed to test the validity of these results using the apparatus depicted in Fig. 5. Light from a single-frequency 6328Å He-Ne laser (200 μ watts) was collimated by a telescope and apertured by a slit. The scattering medium or delay

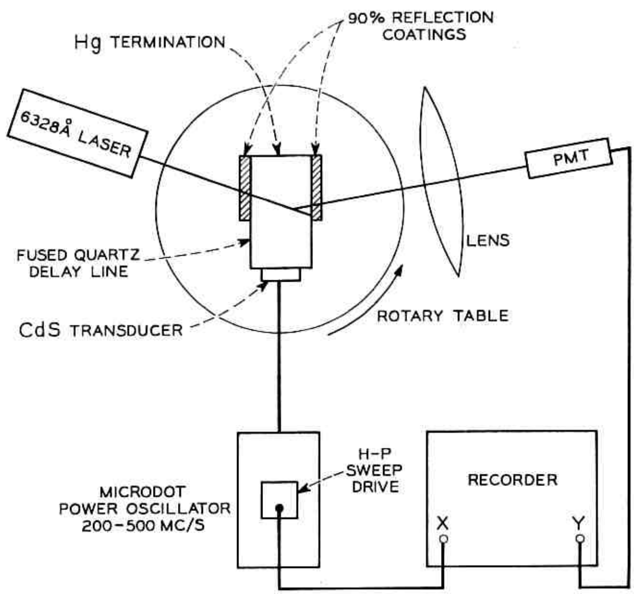

Fig. 5 — Experimental apparatus for studying acoustic scattering in a Fabry-Perot resonator.

line was fused quartz of refractive index $n = 1.457$, width $L = 1.042$ cm, and length $l = 2.5$ cm.

The Fabry-Perot cavity was formed by coating reflecting films over half the length of the delay line which was ground flat and parallel to within one fringe over its entire length. The half nearest the transducer was left without reflecting films to allow comparison of the multiple-pass scattering interaction within the Fabry-Perot cavity with single-pass scattering. The normal Fresnel reflection of the latter half was reduced by the use of quarter-wave matching films in one delay line. A second delay line had no antireflecting films.

The reflectivity of the 5-layer dielectric mirror was determined by transmission measurements to be $R = 0.89$ at 6328Å, resulting in a reflection loss pass of about 11 per cent. A value $R = 0.9$ is close to an optimum compromise between enhancement, transmission loss, and experimental convenience. The dielectric loss of the quartz was expected to be about 0.2 percent per pass yielding $Q_d \approx 50Q_a$ and $T_{ao} = (Q_{L,a}/Q_a)^2 \approx 0.96$. The measured transmission factor for normal incidence was determined to be approximately 0.80. The difference can be related to the lack of perfect parallelism in the opposing faces[10] and to the finite width fact that the light beam was not centered perfectly in the cavity and illustrates the fact that the finite width of the mirrors resulted in walkoff for the larger ring angles. For the case shown, the on-axis mode was not resonant although it could be made resonant by warming the quartz slightly. It should be noted that the actual transmission factors taking only internal losses into account do not exhibit any variation since $\cos \theta_a \approx 1$. An analytical expression for the matching loss is given in Appendix B.

Two delay lines were studied, one with $\Psi <$ one minute of arc and the second with $\Psi \approx 14$ minutes (4×10^{-3} radians). For the first $KL \sin \Psi < \pi/3$ so that scattering could take place over both legs of the round trip. In the second case, $KL \sin \Psi \approx 4\pi$ which allowed interaction on only one leg. The acoustic transducers were evaporated CdS films with an efficiency flat to within 1 dB over the measurement range of 200–450 Mc/s. Using techniques described in Ref. 4, it was determined that the acoustic beam had an essentially uniform intensity distribution over a width of 0.7 cm. The far end of the delay line was terminated with a mercury cell with a reflection loss of 10 dB to inhibit acoustic resonances.

The scattered light was collected and focused onto a Teflon* screen on the face of the photomultiplier; the response was determined to be independent of the scattering angle in the range of interest. The output of the photomultiplier was fed into a phase-sensitive detector, the reference for which was the 1000 cycle/second square wave envelope of the modulated acoustic energy. The output of the phase sensitive detector was displayed on an X-Y recorder. The X-axis drive for the recorder was derived from an angle transducer varying the acoustic frequency. Calibration markers every one or ten megacycles/second were also generated. Absolute acoustic frequency was measured with a counter to within one kilocycle/second. Ring angles were determined to within 10 seconds of arc.

Fig. 7 illustrates typical far-field visual or photographic observations. The Fabry-Perot rings were generated by interposing a Teflon sheet and scattering the incident beam. A second exposure was taken without the Teflon sheet. The main beam was set on the first (fourth) ring and the scattered light appeared on the fourth (first) ring on the opposite side. As expected from (24) the acoustic frequency was essentially, but not exactly, the same for both cases. Closer inspection under high magnification indicated, as expected, that the ring angle for the frequency-

* Trademark of the E.I. du Pont de Nemours, Inc.

shifted, acoustically-scattered light was not quite identical to the equivalent ring angle for the light scattered by the Teflon.

Fig. 8 illustrates an experimental measurement for the case $\Psi = 14'$. The intensity of the scattered light is displayed as a function of acoustic frequency under identical conditions and fixed angle of incidence for single-pass and multiple-pass scattering. The incident beam corresponded to ring #1 and the scattered mode ring numbers are indicated above the peaks. The vertical scale for the multiple-pass scattering was larger by a factor of ten.

The enhancement factor was measured by comparing the single and multiple-pass scattered intensity. Over the frequency range shown, the single-pass scattered intensity variation with frequency was expected to show the main and one upper side lobe of a smooth $(\sin x/x)^2$ variation characteristic of Bragg scattering.[4] The distortion in frequency scale was characteristic of the oscillator frequency drive. These expectations were borne out except for the small bumps which resulted from the low Q Fabry-Perot resonances ($R = 0.035$) associated with the Fresnel reflection at the air-quartz interface. The antireflection coated delay line did not exhibit these bumps.[4] For the case $\Psi = 14'$, the measured

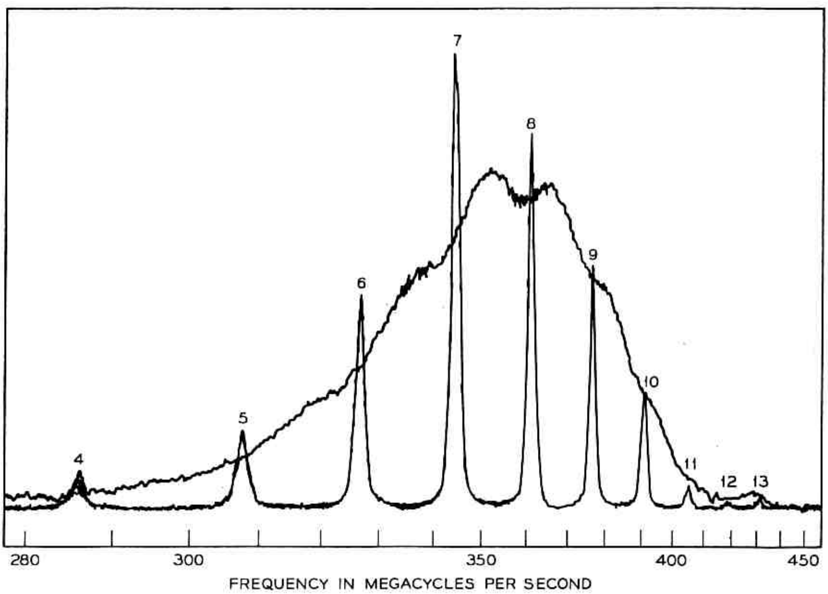

Fig. 8 — Intensity of the scattered light as a function of acoustic frequency for single-pass and multiple-pass scattering. For the latter case, the gain is reduced by a factor of 10.

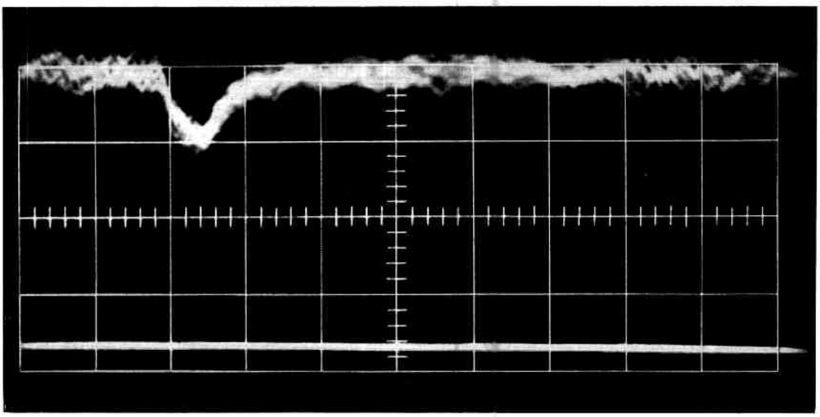

Fig. 9 — Transmitted power as a function of time. The dip corresponds to an acoustic pulse passing through the light beam which removes approximately 25 percent of the power. The acoustic frequency is 265 Mc/s and the time scale is 5 μs/cm.

Modulation depths as large as 25 percent have been observed with about 50 milliwatts of acoustic power. With efficient transducers the required microwave power could be under one watt. These parameters could be improved by the use of more efficient scattering materials than quartz.

V. ACKNOWLEDGMENT

The authors are indebted to N. F. Foster who prepared the transducers; D. L. Perry, who prepared the mirrors; J. R. Wimperis, who fabricated the delay lines, and L. B. Hooker who assisted in the experiments.

APPENDIX A

Derivation of Exponential Decay Rate of Light Within the Cavity

With respect to Fig. 10 it can be seen that when the light travels a distance ds within the cavity, the x-position is changed by an amount $dx = -|\sin\theta_a|\,ds$. In the absence of volume dielectric loss, the light loses energy only at the resonator surface. This loss can be approximated as a volume loss by assuming that the discrete loss upon reflection is distributed uniformly along the path of the light. This approximation is best in the limit $R \to 1$. Thus, if I represents the intensity of the

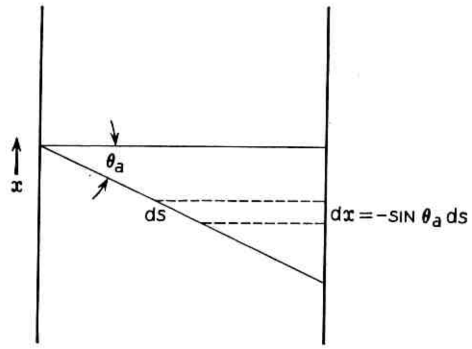

Fig. 10—Cavity geometry for deviation in appendix A.

light, the change in intensity along a path of length ds can be written

$$\begin{aligned} dI &= -I\left(\frac{k_a}{Q_d}\,ds + R^{-\frac{1}{2}}(1-R)\frac{\cos\theta_a\,ds}{L}\right) \\ &= +I\frac{k_a}{|\sin\theta_a|}\left[\frac{1}{Q_d} + \frac{R^{-\frac{1}{2}}(1-R)\cos\theta_a}{k_aL}\right]dx \end{aligned} \tag{42}$$

in which Q_d is the dielectric Q. Noting that $Q_a = k_aL/R^{-\frac{1}{2}}(1-R)\cos\theta_a$ and defining $Q_{L,a}^{-1} = Q_d^{-1} + Q_a^{-1}$, (42) becomes

$$\begin{aligned} dI/dx &= Ik_a/Q_{L,a}\,|\sin\theta_a| \\ &= 2\alpha_a I. \end{aligned} \tag{43}$$

The field amplitude therefore, decays as exp $\alpha_a x$.

The validity of this result can be demonstrated by the following argument. Since the mirror surface has a phase variation given by $-k_a(\sin\theta_a)x$, the far field diffraction pattern of the light transmitted by the mirrors has the form

$$\begin{aligned} T(\theta) &= \left|\int_{-\infty}^{0}\alpha_a dx\,\exp\,[\alpha_a + ik_a(\sin\theta - \sin\theta_a)x]\right|^2 \\ &= \frac{1}{1+[(k_a/\alpha_a)(\sin\theta-\sin\theta_a)]^2}. \end{aligned} \tag{44}$$

Using (43)

$$\begin{aligned} T(\theta) &= \frac{1}{1+[2Q_{L,a}\sin\theta_a(\sin\theta-\sin\theta_a)]^2} \\ &\approx \frac{1}{1+[2Q_{L,a}\cos\theta_a(\cos\theta-\cos\theta_a)]^2}. \end{aligned} \tag{45}$$

enhancement factors and bandwidths agreed to about 5 percent which was better than could be expected.

The most detailed measurements were made for the case $\Psi < 1'$. Verification of (24) for the frequency of the interaction was obtained by determining $\Omega_{a,a'}$ and θ_a and writing (24) as

$$v = \frac{\lambda_0(\Omega_{a,a'}/2\pi)\cos\Psi}{n[|\sin\theta_a| + |\sin\theta_{a'}|(1 \pm \Omega_{a,a'}/\omega_a)\cot^2\theta_{a'})]} \tag{38}$$

in which λ_0 is the vacuum wavelength and n the index of refraction of the quartz, and computing v from the measured frequencies and mode angles. The angles θ_a were sufficiently small (of order 10^{-2} radians) that $n \sin\theta_a$ could be taken equal to the externally measured ring angles and λ_0 to the air wavelength. The measured ring angles were consistent with those computed for the cavity. The data yielded a value $v = 5.940 \pm 0.005 \times 10^5$ cm/sec, consistent with the uncertainty in measuring θ_a. Measurements in the single-pass scattering region were also made by determining the Bragg angle Θ for scattering to $\omega \pm \Omega$ defined by

$$\sin\Theta_\pm = v/c' \mp \tfrac{1}{2}K/k' \tag{39}$$

which yields

$$\Theta_- - \Theta_+ \approx \sin\Theta_- - \sin\Theta_+ = K/k'. \tag{40}$$

The phase velocity is determined by

$$v = (\Omega/2\pi)\lambda_0/\Delta\theta \tag{41}$$

in which $\Delta\theta$ is the externally determined difference between the two Bragg angles. The phase velocity was determined by this technique to be 5.940×10^5 cm/sec with the same spread in measured values, in agreement with the optical Fabry-Perot measurements. The phase velocity was also determined by setting the incident light at the Bragg angle in the single-pass region and using the intensity of the scattered light as a measure of the acoustic intensity while the acoustic frequency was varied through a range of 2 megacycles/second. The acoustic Fabry-Perot resonances were emphasized by removing the mercury from the far end. Using the relation

$$\Delta\Omega/2\pi = v/2l$$

in which l is the length of the fused quartz bar and $\Delta\Omega/2\pi$ is the spacing of the resonances yielded a value $v = (5.950 \pm 0.003) \times 10^5$ cm/second. While the difference here is small, it falls outside the range of experimental uncertainty. The difference has been ascribed tentatively to pulling of the acoustic resonances by the transducer. The values of v quoted in the literature[11,12] fall in the range $5.90 - 5.96 \times 10^5$ cm/second.

Comparison with theory of experimentally determined values of enhancement and bandwidth is hampered slightly by an uncertainty in the value of $(1 - R)$ which could be as large as ± 10 percent. On the other hand, the product

$$(\Delta\Omega/2\pi) g_{\text{opt}}^{\frac{1}{2}}$$

using (33), (37), and (36) is independent of R. Preliminary measurements for the case $\Psi = 14'$ were very encouraging. Since the case $\Psi < 1'$ was believed to be more interesting most of the measurements were made here. To simplify interpretation, the measurements were restricted to scattering from a given ring to the same ring on the other side of normal incidence $(a \rightarrow a)$. Some of these measurements are summarized in Table I. The calculations were performed taking $T_{ao} = 0.96$, however, correcting g_{opt} for the matching loss as experimentally determined. In all cases both $\Delta\Omega/2\pi$ and $g_{\text{opt}}^{\frac{1}{2}}$ were larger than calculated, the product being as much as 80 percent larger than expected. Values smaller than expected could possibly be attributed to an imperfect Fabry-Perot cavity. No explanation for the larger values can be offered at this time.

Under optimum conditions, the modulation depth was approximately 25 percent as shown in Fig. 9. The acoustic power was estimated to be about 50 milliwatts.

IV. CONCLUSION

Experiments illustrating acoustic scattering in a Fabry-Perot resonator have been described and compared with a coupled-mode theory. The acoustic frequencies for resonant scattering agree with theory to within the experimental uncertainty of 1:500. The agreement on enhancement factors and bandwidth is within 80 percent but well outside experimental uncertainty; the measured values are on the large side.

TABLE I

Ring	$(\Omega_{a,a}/2\pi)$ (Mc/s)	g_{ex}	$(\Delta\Omega/2\pi)_{ex}$ Mc/s)	$(g^{\frac{1}{2}}\Delta\Omega/\pi)_{ex}$ (Mc/s)	$(g^{\frac{1}{2}}\Delta\Omega/2\pi)_{calc.}$ (Mc/s)
2	220	73.3	3.1	26.5	17
3	283	47	3.27	22.5	12.5
4	333	49	1.94	13.6	10.5

Note that when there is no internal loss, $Q_{L,a} = Q_a = k_a L R^{\frac{1}{2}}/(1 - R)$ $\cos \theta_a$ and

$$T(\theta) \approx \frac{1}{1 + [4R/(1 - R)^2][k_a L(\cos \theta - \cos \theta_a)]^2} \tag{46}$$

which corresponds very closely to the Haidinger fringe formula in the limit $R \to 1$. Thus, the approximate expression for the optical decay (43) is most valid in this limit.

It can be shown that the difference between the Lorentzian form of (46) and the true Haidinger fringe formula arises from the approximation in which the discrete reflection is replaced by a fictitious volume loss; the stepwise decay of the light intensity becomes an exponential decay. This approximation is, of course, best when the steps become vanishingly small, i.e., when $R \to 1$. A demonstration of the more precise result is straightforward but lengthy and will not be included here.

APPENDIX B

Derivation of Cavity Transmission for Off-Axis Modes

Assuming that the cavity transmission for a given mode may be approximated by $T(\theta)$ given in (46) and the energy distribution of the incident beam by $F(\theta)$ then the transmission factor may be written

$$T = \frac{\int_{-\pi/2}^{+\pi/2} F(\theta)T(\theta)d\theta}{\int_{-\pi/2}^{+\pi/2} F(\theta)d\theta}. \tag{47}$$

For an incident beam with rectangular cross section of width W incident at angle θ_a,

$$F(\theta) = \left[\frac{\sin \frac{1}{2}kW(\sin \theta - \sin \theta_a)}{\frac{1}{2}kW(\sin \theta - \sin \theta_a)}\right]^2 \tag{48}$$

and

$$T \approx \frac{\int_{-\infty}^{+\infty} [\sin^2 x/x^2(1 + a^2x^2)]dx}{\int_{-\infty}^{+\infty} [\sin^2 x/x^2]dx} \tag{49}$$

in which

$$a^2 = \frac{16R}{(1 - R)^2}\left[\frac{L \tan \theta_a}{W}\right]^2.$$

The integral in (49) can be evaluated by a straight-forward contour integration to yield

$$T = 1 - ae^{-a^{-1}} \sinh a^{-1}. \tag{50}$$

In the limit $(L \tan \theta_a/W) \propto a \rightarrow 0$, $T = 1$, as would be expected, since the angular spread in the incident beam is small compared to the angular spread of the cavity mode. In the opposite limit $(L \tan \theta_a/W) \propto a \rightarrow \infty$, $T \approx a^{-1}$ and the transmission factor for successive cavity modes of increasing angle θ_a varies as cot θ_a.

REFERENCES

1. Born, M., and Wolf, E., *Principles of Optics*, Pergamon Press, New York, 1959, Chap. XII.
2. Benedek, G. B., Lastovka, J. B., Fritsch, K., and Greytak, T., Brillouin Scattering in Liquids and Solids Using Low Power Lasers, J. Opt. Soc. Amer., *54*, October, 1964.
3. Chiao, R. Y. and Stoicheff, B. P., Brillouin Scattering in Liquids Excited by the He-Ne Maser, J. Opt. Soc. Amer., *54*, 1964, p. 1286.
4. Cohen, M. G. and Gordon, E. I., Acoustic Beam Probing Using Optical Techniques, B.S.T.J., *44*, 1965, p. 693.
5. Gordon, E. I. and Cohen, M. G., Electro-optic Diffraction Gratings for Light Beam Modulation and Diffraction, IEEE J. Quantum Elect., *1*, August, 1965
6. Cohen M. G. and Gordon, E. I., Electro-optic [$KTa_xNb_{1-x}O_3$ (KTN)] Gratings for Light Beam Modulation and Deflection, Appl. Phys. Let., *5*, 1964, p. 181.
7. Gordon, E. I. and Rigden, J. D., The Fabry-Perot Electro-optic Modulator, B.S.T.J., *42*, 1963, p. 155.
8. Jenkins, F. A. and White, H. E., *Fundamentals of Optics*, McGraw-Hill Book Company, Inc., New York, 1937, p. 87.
9. Gordon, E. I., Figure of Merit for Acousto-Optical Deflection and Modulation Devices, IEEE J. Quantum Elect., *2*, (Correspondence), May, 1966, p. 104.
10. Del Piano, V. N., Jr. and Quesada, A. F., Transmission Characteristics of Fabry-Perot Interferometers and a Related Electro-optic Modulator, Appl. Opt., *4*, November, 1965, pp. 1386–1390.
11. Mason, W. P., *Physical Acoustics and the Properties of Solids*, D. Van Nostrand Co., Inc., 1958, p. 17.
12. *Handbook of Chemistry and Physics*, Chemical Rubber Co., Cleveland, Ohio, 1964.

of the incident light beam which resulted in an angular spread comparable to or larger than the angular width of the Fabry-Perot modes. The measured transmission factors for off-axis modes were smaller than the value for the on-axis mode since the angular width decreases with increasing mode angle. This is illustrated in the photograph of Fig. 6 which displays the transmission factor of the cavity as a function of angle of the incident light relative to the surface normal. The upper line is unity transmission. The measured transmission factors became increasingly smaller for increasing mode angle indicating that as the angular spread of the rings became increasingly smaller, less of the incident light was matched into the cavity mode. This situation, while normally undesirable and easily avoidable, was experimentally convenient since it guaranteed that the dominant feature of the light distribution in the incident beam mode was an exponential decay as postulated in Section II. The incident beam mode usually corresponded to a relatively small angle and the matching loss was not excessively large. The lack of perfect symmetry about normal incidence arose from the

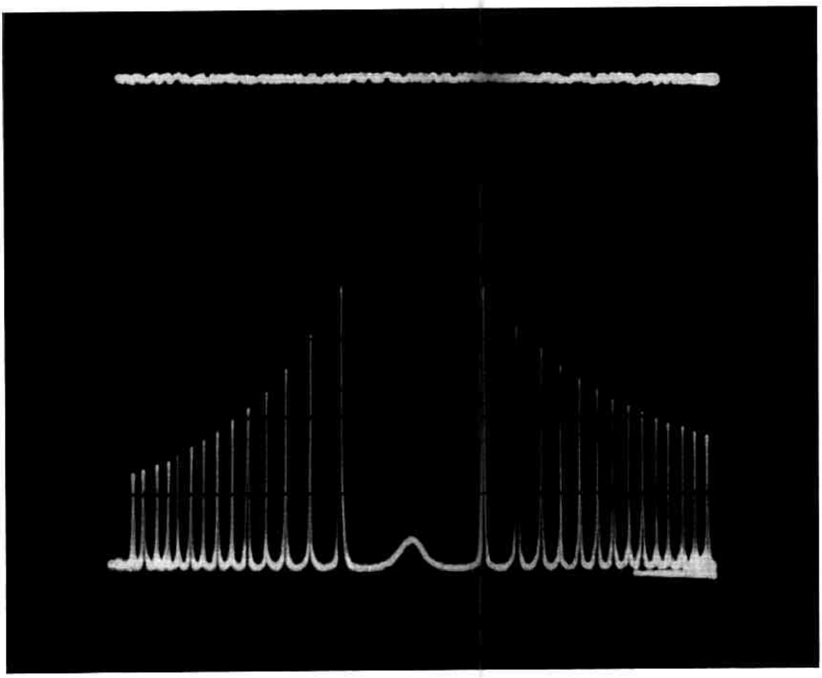

Fig. 6 — Measured Fabry-Perot cavity transmission factors as a function of angle of the incident light relative to the surface normal.

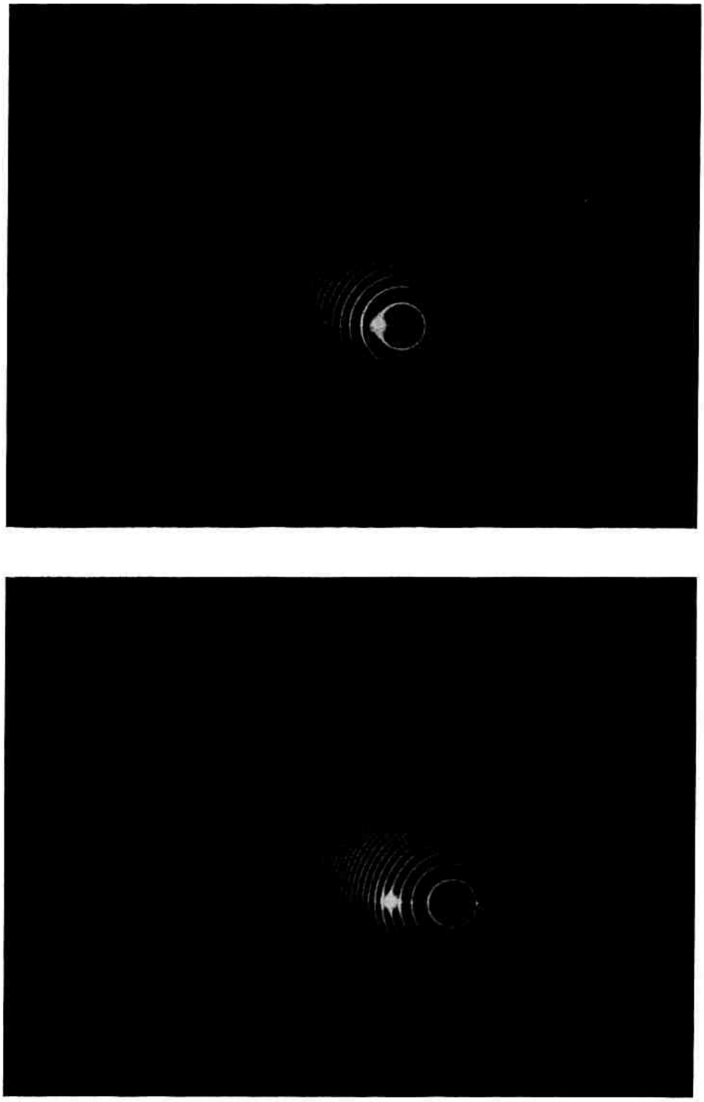

Fig. 7 — Photographs of typical far-field observations of scattering in the Fabry-Perot resonator. The angles of the incident and scattered beams are interchanged in the two photographs.

A Bound on the Reliability of Block Coding with Feedback

By J. E. SAVAGE

(Manuscript received March 25, 1966)

A conceptually simple, block-coding feedback strategy which applies to all time-discrete, memoryless channels is introduced and examined. This strategy provides the first conclusive evidence of an improvement at nonzero rates in the reliability of block coding with feedback on the additive Gaussian noise channel. This result was previously observed by Berlekamp for the binary symmetric channel.

I. INTRODUCTION

Coding with feedback has been considered by a number of authors.[1,2,3,4,5,6] principally because of the advantage it is expected to enjoy in rate, reliability, and equipment costs over coding without feedback. Some early feedback results were obtained by Shannon, who showed that channel capacity of a discrete memoryless channel (DMC) cannot be increased using feedback[1] and established that the sphere-packing (lower bound[7,8]) to the probability of error without feedback applies to *block coding* with feedback on the DMC uniform at the input (the sets of transition probabilities from each channel input letter are identical except for permutations[7]). He also conjectured that a sphere-packing bound applies to *block coding* with feedback on all discrete, memoryless channels.[2] (A conjecture supported recently by Berlekamp.[9])

It has been shown that the exponent on the sphere-packing bound agrees with the no-feedback, random-code exponent at rates above the critical rate R_{crit},[7,10] so that feedback cannot increase the reliability of block coding at rates greater than R_{crit}. Below R_{crit}, Berlekamp first showed that feedback will improve the reliability of block coding.[3] He showed that the zero rate exponent of the probability of error with block codes on the binary symmetric channel (BSC) and certain other binary channels having particular symmetries is larger with feedback than the best no-feedback exponent at zero rate. Also, he showed that the error-correction capability of block codes on the BSC at rates between zero and R_{crit} is improved with feedback.

Variable-block-length coding strategies, that is strategies where the code length is controlled through the feedback channel, have been proposed and in contrast with the block coding strategies have been shown to operate with a greater reliability than that given by the sphere-packing bound.[4,5] Hence, the best variable-length feedback strategy is superior to the best block-coding feedback strategy. The block-coding strategies, however, are interesting since any improvement in coding reliability observed with them can be ascribed directly to the effect of feedback on the choice of codewords representing messages. This is not true of the variable-length strategies since, in this case, some unknown fraction of the improvement in reliability is attributable to the variation of the code length with the level of channel noise.

In this paper, we shall be concerned only with *block-coding* with feedback and in particular with one particular block-coding strategy. This is a strategy which makes efficient but not complete use of the feedback channel. (This will be seen from the results.) It is recommended, however, by its simplicity, its easy evaluation in terms of known, (no-feedback) bounds, by its application to a large class of channels including all time-discrete, memoryless channels, by the substantial improvements it shows for many channels over the no-feedback, random-code exponents and the improvement it shows for some channels over the best upper bounds on no-feedback exponents. In particular, we see an improvement over the random-code exponents over a middle range of rates on the BSC with crossover probability less than 10^{-7} and on the average-power-limited, additive Gaussian noise channel (AGC) with signal-to-noise ratio (S/N) greater than 11.5 dB. This is interesting, since some believe that the random-code exponents are the best, no-feedback, block-coding exponents. Also, we see an improvement, again over a middle range of rates, over the Wyner upper bound[11] to the no-feedback exponent on the AGC with S/N $\geqq$ 22 dB.

It is clear that our strategy does not make full use of the feedback channel since the improvements noted apply only to cleaner channels and then only over a middle range of rates. Also, it can be shown that our strategy has an exponent which is smaller than the exponent implied by Berlekamp's results for the BSC.[3] It should be noted, however, that the BSC is the only channel for which it has been previously shown that feedback can improve the reliability of block coding.

II. THE CODING STRATEGY

The feedback, block-coding strategy which we present was suggested by arguments used by Shannon and Gallager to *underbound* the probability of error with block coding.[12] Our strategy, however, leads to an overbound to the probability of error with the best feedback strategy since it does not make the most efficient use of the return channel. We will describe this strategy without referring directly to the input and output alphabets of the forward channel. We assume only that the forward channel is time-discrete. With respect to the return channel, we assume that it is noiseless and of large but finite capacity. The forward and reverse channels are allowed to have a total delay of D channel symbols.

Our feedback, block-coding strategy is a 2-step procedure. For each of the two steps a block code is chosen and used without feedback. In the first step, a codeword from a code of M codewords each having length N_1, is chosen to transmit one of the M messages generated by the source. In the second step, a code of L codewords* where each word has length N_2 is used. Here N_1 and N_2 are chosen so that $N_1 + N_2 + D = N$, the number of channel symbol intervals allowed for the transmission of one of the M source messages.

The decoder receives a noisy version of the codeword chosen from the first code and he then makes a list of the L messages which are most likely given the received signal.† We assume that all messages are equally likely, *a priori*, so that the L messages on the list are those which have the largest likelihood probability. That is, if $p(\mathbf{v}_1 \mid m)$ is the probability (or probability density, if the channel output alphabet is continuous) of receiving the N_1 channel letters represented with $\mathbf{v}_1$ when message m is transmitted, then messages $m_1, m_2, \cdots, m_L$ are on the list if

$$p(\mathbf{v}_1 \mid m_i) \geqq p(\mathbf{v}_1 \mid m') \qquad \text{all } m' \neq m_i, \qquad 1 \leqq i \leqq L. \quad (1)$$

Once the list has been formed, the list and the order in which messages appear on the list is sent over the reverse channel to the transmitter. Since D time intervals will elapse before the list reaches the transmitter, the transmitter is ready to begin the second of the two steps after $N_1 + D$ intervals. In the second step, the transmitter chooses a codeword from the second code of length N_2 to indicate which message on

* L is arbitrary here but is fixed in later discussion.
† This decoding procedure is called "list decoding."[8,13]

the list of the L messages is the source output. If the source output is not on the list, the first codeword is transmitted. The received sequence is decoded using the max-likelihood decoding rule.

III. EVALUATION OF THE STRATEGY

Our coding strategy which first establishes a list of most probable transmitted messages and then resolves the ambiguity in the list can lead to a decoding error in either of two ways. First, the message delivered by the source may not be on the list because of excessive channel noise in the first N_1 transmissions or, second, it may be on the list but the list may be decoded in error during the last N_2 transmissions.

Now, the probability of a decoding error with the best feedback strategy, $P_{ef}(N,M1)$, is less than or equal to $P_f(N,M,1)$, the probability of error, with the feedback strategy given above. This, in turn, is less than or equal to the sum of the probability of list decoding error with the best list code (and no feedback), $P_e(N_1, M, L)$, and the probability of a max-likelihood decoding error with the best max-likelihood code, $P_e(N_2, L, 1)$. Thus, we have

$$P_F(N,M,1) \leqq P_f(N,M,1) \leqq P_e(N_1, M, L) + P_e(N_2, L, 1) \quad (2)$$

where

$$N = N_1 + N_2 + D \quad (3)$$

and D is the round-tip delay.

While our feedback strategy can be evaluated for any forward channel for which bounds to $P_e(N_1, M, L)$ and $P_e(N_2, L, 1)$ are known, we restrict our attention here to the discrete memoryless channel and to the time-discrete, average-power-limited, additive Gaussian noise channel. Bounds to these error probabilities for these two channels can be found in several places.[7,10,14,15] However, for easy reference we shall refer to Gallager.[10] Gallager does not bound $P_e(N_1, M, L)$ in his paper but the changes necessary in his analysis to bound it are relatively easy to effect and are outlined in the Appendix. These changes are due to unpublished results by Gallager* and are such that the random-code bound to $P_e(N_1, M, L)$ has the same form as Gallager's bound[10] to $P_e(N_1, M, 1)$ except that his parameter ρ is allowed to range between 0 and L rather than between 0 and 1.

The bounds to $P_e(N_1, M, L)$ and $P_e(N_2, L, 1)$ are given below where $R_1 = (\log_2 M)/N_1$, $R_2 = (\log_2 L)/N_2$ and $O_i(N_i)$, $1 \leqq i \leqq 2$, are quan-

* See also Ref. 16.

tities which approach zero faster than $1/N_i$:

$$P_e(N_1, M, L) \leqq 2^{-N_1[E_L(R_1)+O_1(N_1)]} \tag{4}$$

$$P_e(N_2, L, 1) \leqq 2^{-N_2[E_x(R_2)+O_2(N_2)]}. \tag{5}$$

Here $E_L(R_1)$ and $E_x(R_2)$ are the random-code bound and expurgated random-code bound, respectively, to the exponents on the two error probabilities. The formal statement of these two exponents for the DMC and the AGC is somewhat long. Consequently, we present Table I which lists the location of the two exponents by equation number in Ref. 10. We note again that $E_L(R_1)$ has the same form as $E_1(R_1)$, which is the random-code exponent given by Gallager, except that $0 \leqq \rho \leqq L$.

Equations (4) and (5) are used to bound (2). The block lengths

TABLE I — LOCATION OF EXPONENTS IN REF. 10.

	DMC	AGC
$E_L(R_1)$	21,22	125,127,128
$E_x(R_2)$	86,87	133

N_1 and N_2 of the two codes are chosen before transmission to approximately optimize the bounds (4) and (5). That is, we set

$$N_1E_L(R_1) = N_2E_x(R_2). \tag{6}$$

Since $N_2 = (N - D) - N_1$ we have

$$N_1 = \frac{(N - D)E_x(R_2)}{E_x(R_2) + E_L(R_1)}. \tag{7}$$

Thus, the exponent $N_1E_L(R_1)$ becomes

$$N_1E_L(R_1) = \frac{(N - D)E_x(R_2)E_L(R_1)}{E_x(R_2) + E_L(R_1)}. \tag{8}$$

The signaling rate of the code is defined as $R = (\log_2 M)/N$ so that

$$R = \frac{N_1}{N} R_1 = \frac{\left(1 - \frac{D}{N}\right) R_1E_x(R_2)}{E_x(R_2) + E_L(R_1)}. \tag{9}$$

The exponent to the probability of error with our feedback strategy, $E_f(R)$, is defined as

$$E_f(R) = \lim_{N\to\infty} -\frac{\log_2 P_f}{N} \tag{10}$$

so that if we assume that $D/N \to 0$ *as* N *increases* and if L *is independent of* N $(R_2 \to 0)$, then

$$E_f(R) = \frac{E_x(0)E_L(R_1)}{E_x(0) + E_L(R_1)} \tag{11}$$

where

$$R = \frac{R_1 E_x(0)}{E_x(0) + E_L(R_1)} \cdot \tag{12}$$

Hence, $E_f(R)$ and R are parametrized by the rate R_1.

Let us now consider the list size L and show that it should be made independent of N. The list size appears in (11) and (12) through the list decoding exponent $E_L(R_1)$. As mentioned above, this exponent is parametrized with a parameter ρ, $0 \leqq \rho \leqq L$. Associated with $E_L(R_1)$ is a rate R_L^* such that if $R_1 < R_L^*$ then $E_L(R_1)$ is a straight line of slope $-L$. Also, if $R_1 > R_L^*$ then $E_L(R_1)$ has slope $-\rho$, $\rho < L$. Thus, $E_L(R_1)$ is largest for fixed R_1 if L is such that $R_1 > R_L^*$. For any R_1, the L satisfying this inequality is fixed and finite. Now, examination of (11) and (12) will show that $E_f(R)$ vs R is largest when $E_L(R_1)$ is largest so that $E_f(R)$ is maximized over L with a value of L which is independent of N. We choose to use that L for which $R_1 > R_L^*$ so that the arguments made in the previous paragraph hold.

It can be shown[7] that the exponent $E_L(R_1)$ is equal to the sphere-packing bound $E_{sp}(R_1)$ for $R_1 \geqq R_L^*$. Thus, we have, as our final result, the achievable (lower) bound, $E_f(R)$, to the *largest obtainable exponent with feedback*, $E_F(R)$, given below:

$$E_F(R) \geqq E_f(R) = \frac{E_x(0)E_{sp}(R_1)}{E_x(0) + E_{sp}(R_1)} \tag{13}$$

where

$$R = \frac{R_1 E_x(0)}{E_x(0) + E_{sp}(R_1)} \cdot \tag{14}$$

A simple construction for $E_f(R)$ from $E_x(0)$ and the sphere-packing bound is shown in Fig. 1.

IV. SOME EXAMPLES

The exponent $E_f(R)$ on the BSC is significantly smaller than the exponent implied by Berlekamp's results and for this reason is not shown.

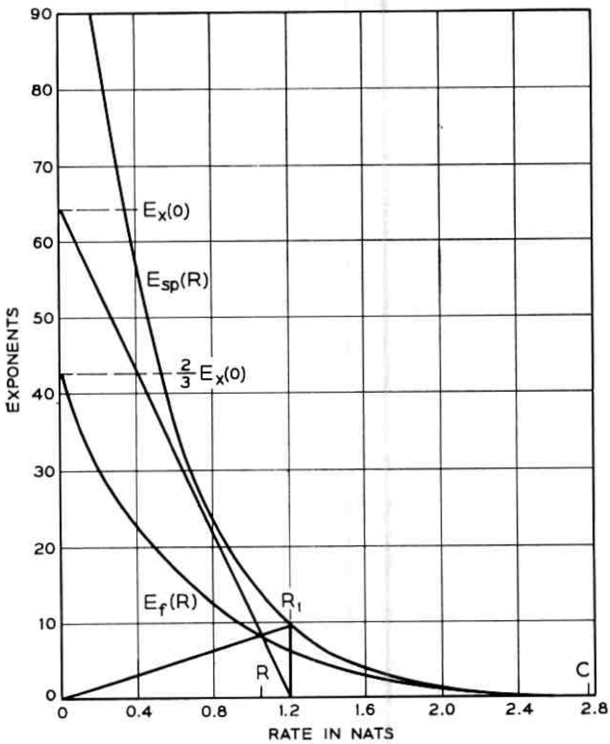

Fig. 1 — Construction for $E_f\,(R)$ on AGC, S/N = 256.

A computer study of $E_f(R)$, however, shows that $E_f(R)$ is larger than the random code bounds with and without expurgation over a middle range of rates when the crossover probability $p \geqq 10^{-7}$. It is found that this range of rates increases with decreasing p.

The exponent $E_f(R)$ is shown in Figs. 2 and 3 for the time-discrete, average-power-limited, additive Gaussian noise channel with power signal-to-noise ratios of 64 (18.1 dB) and 256 (24.1 dB), respectively. In the first case, an improvement is seen over both the expurgated and unexpurgated random code exponents for a very large range of rates, namely, $0.15 \leqq R \leqq 1.6$, in nats, and in the second case $E_f(R)$ is larger than the Wyner *upper bound*[11] to the exponent without feedback, $E_w(R)$, for a substantial range of rates, namely $0.90 \leqq R \leqq 1.80$. Computer calculations have shown that if S/N < 160 (22 db), then $E_f(R) < E_w(R)$ and if S/N < 14 (11.5 dB), then $E_f(R) \leqq E_{rc}(R)$, which is the random-code exponent without feedback.

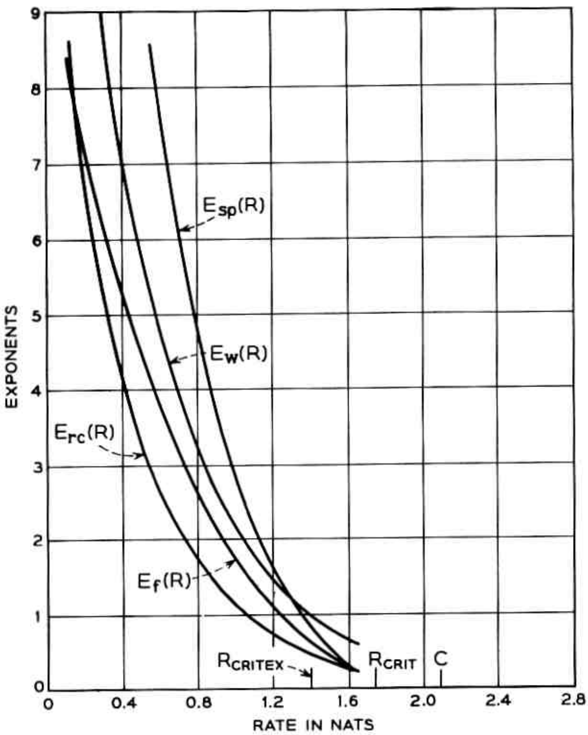

Fig. 2 — Exponents on AGC, S/N = 64.

V. CONCLUSIONS

We have introduced and examined a conceptually simple block-coding, feedback strategy. Using this strategy as an example of block-coding feedback strategies, we have found the first conclusive evidence of an improvement in the reliability of block coding with feedback at nonzero information rates on the additive Gaussian noise channel. This improvement is measured with the exponent on the probability of error and we have shown that an exponent larger than the random-code (lower bound) exponent can be obtained on many channels such as the relatively clean BSC and AGC and that on at least one channel, the AGC with S/N $\geqq$ 22 dB, an exponent which is superior to the best no-feedback exponent can be achieved. These results have been shown to hold when there is a nonzero channel delay as long as that delay does not grow as fast as linearly with block length.

The gain in reliability seen above can be translated into reduced

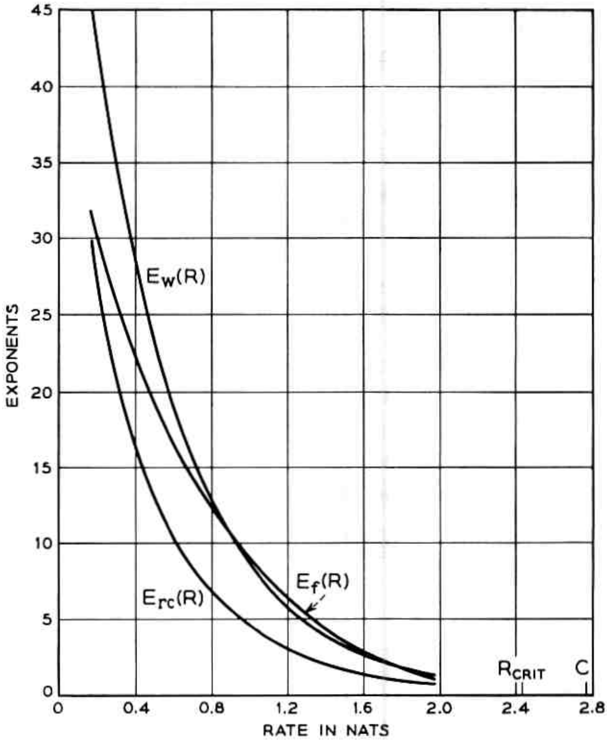

Fig. 3 — Exponents on AGC, S/N = 256.

equipment costs or increased rate. We have made our comparison of coding with and without feedback on the basis of fixed rate and block length.

It is important to emphasize that the improvement in the performance of block coding using feedback is due entirely to the fact that the codewords in the code are allowed to change with the channel noise. In our example, the list formed after the first step changes with the level of the channel noise so that the association of codewords in the second code with messages on the list becomes channel dependent.

VI. ACKNOWLEDGMENT

The author gratefully acknowledges many helpful conversations with P. M. Ebert during this work. He also acknowledges his help in the calculation of the Gaussian channel exponents.

APPENDIX

Equation (1) states the rule for choosing messages $m_1, m_2, \cdots, m_L$ to be placed on the list when list decoding of list size L is used and the channel sequence $\mathbf{v}_1$ is received. An error is made if the transmitted message m, say, is not on this list. Using an argument similar to that given in Ref. 10, (2) through (6), the probability of error with list decoding when message m is received can be stated formally for a particular code as

$$P_{em} = \sum_{\mathbf{v}_1 \in V_{N_1}} P(\mathbf{v}_1 \mid \mathbf{x}_m)\Phi_m(\mathbf{v}_1).$$

Here V_{N_1} is the set of channel sequences $\{\mathbf{v}_1\}$ of length N_1, $\mathbf{x}_m$ is the mth code word in the code and $\Phi_m(\mathbf{v}_1)$ is a characteristic function which is one if $\mathbf{v}_1$ results in a list which does not include m and is zero otherwise. Gallager, in unpublished work, has shown that the characteristic function $\Phi_m(\mathbf{v}_1)$ may be overbounded by the following

$$\Phi_m(\mathbf{v}_1) \leq \left\{ \sum_{m_1'} \cdots \sum_{\substack{m_L' \\ m_i' \neq m \\ 1 \leq i \leq L}} \frac{1}{L!} \frac{P(\mathbf{v}_1 \mid \mathbf{x}_{m_1'})^{1/1+\rho} \cdots P(\mathbf{v}_1 \mid \mathbf{x}_{m_L'})^{1/1+\rho}}{P(v_1 \mid \mathbf{x}_m)^{1/1+\rho} \cdots P(\mathbf{v}_1 \mid \mathbf{x}_m)^{1/1+\rho}} \right\}^{\rho/L}$$

where $\rho \geqq 0$. Carrying out the random code arguments as given in Ref. 10 with $0 \leqq \rho \leqq L$, we have the desired result.

REFERENCES

1. Shannon, C. E., Zero Error Capacity of a Noisy Channel, IRE Trans, *IT-2,* September, 1956, pp. 8–19.
2. Shannon, C. E., unpublished seminar notes numbered 1p–3p, 1956.
3. Berlekamp, E. R., Block Coding with Noiseless Feedback, M.I.T. Thesis, August, 1964.
4. Horstein, M., Sequential Transmission Using Noiseless Feedback, IRE Trans, *IT-9,* 1963, pp. 136–143.
5. Weldon, E. J., Asymptotic Error Coding Bounds for the Binary Symmetric Channel with Feedback, Ph.D. Thesis, Dept. of Electrical Engineering, University of Florida, April, 1963.
6. Chang, S. S. L., Theory of Information Feedback Systems, IRE Trans, *IT-11,* 1956, pp. 29–40.
7. Fano, R. M., *Transmission of Information,* M.I.T. Press and John Wiley and Sons, New York, 1961.
8. Wozencraft, J. M., List Decoding, Quarterly Progress Report No. 18, Research Laboratory of Electronics, M.I.T., January 15, 1958.
9. Berlekamp, E. R., private communication.
10. Gallager, R. G., A Simple Derivation of the Coding Theorem and Some Applications, IEEE Trans, *IT-11,* January, 1965, pp. 3–17.
11. Wyner, A. D., An Improved Error Bound for Gaussian Channels, B.S.T.J., *63,* November, 1964, pp. 3070–3074.
12. Shannon, C. E., Gallager, R. G., and Berlekamp, E. R., Lower Bounds to Error Probability for Coding on Discrete Memoryless Channels, submitted to Information and Control.

13. Elias, P., List Decoding for Noisy Channels, Technical Report No. 335, Research Laboratory of Electronics, M.I.T., September 20, 1957.
14. Elias, P., Coding for Two Noisy Channels, *Information Theory* (C. Cherry, Editor), Butterworth, London, 1956.
15. Shannon, C. E., Probability of Error for Optimal Codes in a Gaussian Channel, B.S.T.J., *38*, May, 1959, pp. 611–656.
16. Ebert, P. M., Error Bounds for Parallel Communication Channels, M.I.T. Thesis, September, 1965.
17. Gallager, R. G., unpublished notes.

Error Probabilities in Data System Pulse Regenerator with DC Restoration

By P. L. ZADOR

(Manuscript received March 17, 1966)

Consider a noisy channel which acts as a high-pass filter on the pulses used for transmitting digital data in binary form. To combat the degradation of information in the channel, the pulses are detected and regenerated at certain points with the aid of a binary pulse regenerator with dc restoration. This device achieves complete restoration in the absence of noise.

In this paper, we give a procedure for evaluating the limiting probabilities (after lengthy operation) of error patterns for a single dc restorer in the presence of independent, additive noise. The procedure is based on the observation that for the particular restorer in question, the effective noise in the restorer is the sum of the present noise and the accumulated noise. The latter may be described by a Markovian process.

I. INTRODUCTION

In this paper, we consider a binary pulse regenerator with dc restoration (Fig. 1). This system has the property that in the absence of noise it functions error-free. To evaluate the performance of the system, one would like to know the probability of occurrence for an arbitrary error pattern when noise is introduced. This is highly desirable especially if the information coding places high penalty on certain error patterns. It will be shown in this paper that the limiting probability of any error burst can be computed by an iterative procedure. The mathematical theory justifying the validity of the procedure is somewhat involved and will be given in a separate paper on Random Walk in Compact Metric Space.[1]

II. THE SYSTEM

Our considerations apply to the data transmission system represented by the block diagram given in Fig. 1. The input sequence d_k of random ± 1 impulses goes through a high-pass filter. The output of this filter is

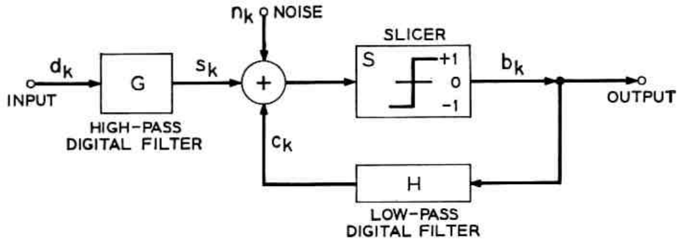

Fig. 1 — Binary pulse regenerator with low-frequency restoration.

sampled synchronously with the impulse train input to yield s_k. It is assumed that the output is contaminated by independent noise n_k. A slicer triggering ± 1 impulses b_k is used for regeneration of pulses. There is a low-pass filter in the feedback loop followed by a sample r which is also synchronized with the input sequence. In the block diagram G and H represent the filters followed by the samplers.

G is a discrete high-pass and H is a discrete low-pass linear filter. S is assumed to be an ideal slicer discriminating between positive and negative voltage levels. Thus, we have the system equations

$$s_k = \sum_{i=0}^{k} g_{k-i} d_i \qquad k = 0,1, \cdots \tag{1}$$

$$c_k = \sum_{i=0}^{k} h_{k-i} b_i \qquad k = 0,1, \cdots \tag{2}$$

and

$$b_k = \operatorname{sgn} \{n_k + c_k + s_k\} \tag{3}$$

where $\operatorname{sgn} x = 1$ if $x \geqq 0$ and $\operatorname{sgn} x = -1$ if $x < 0$, and $g_i(h_i)$ is the impulse response of the filter $G(H)$ at time i. As to the filters G and H, it is assumed that

(i) $\qquad h_o = 0, \qquad g_0 > 0$

(ii) $\qquad h_i + g_i = 0 \quad \text{if} \quad i \neq 0$

(iii) $\qquad 0 > g_i = r g_{i-1} \quad \text{for} \quad i \geqq 2$

where $0 < r < 1$. The interpretation of $(i$–$iii)$ is as follows: (i) There is unit time delay in the feedback loop. (ii) The filter H is chosen to cancel the tails of the impulse response of G, a condition allowing the error free operation of the system in the absence of noise. (iii) The channel has exponentially decaying impulse response. It is this last property that will allow us to describe the cumulative error by a Markov process.

The input and noise sequences are regarded as samples from two sequences of completely independent random variables (r.v.). We assume that $d_k = 1$ or $d_k = -1$ with fixed but arbitrary probabilities p and q. The r.v.'s n_k representing the noise have a fixed and, for the time being, arbitrary distribution function (d.f.) $N(x)$. Typically, $N(x)$ is the normal d.f., i.e.,

$$N(x) = \frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^{x} \exp\,(x^2/2\sigma^2 dx). \tag{4}$$

We say that the kth output b_k is in error if $b_k \neq d_k$.

III. EFFECTIVE NOISE

When noise is absent from the system ($n_0 = 0,\ n_1 = 0,\ \cdots$) then $b_k = d_k$ for $k = 0, 1, \cdots$. Clearly, $b_0 = \mathrm{d}_0$. Suppose that $b_i = d_i$ for $i = 0, \cdots, k-1$. It follows from (1), (2) and (*ii*) that in this case $s_k + c_k = d_k$. Hence, $b_k = d_k$ due to (3). The system is error-free in the absence of noise. To clarify the effect of noise, notice that

$$\begin{aligned} s_k + c_k = \sum_{\substack{i=0 \\ b_i = d_i}}^{k-1} (h_{k-i}b_i + g_{k-i}d_i) \\ + \sum_{\substack{i=0 \\ b_i \neq d_i}}^{k-1} (h_{k-i}b_i + g_{k-i}d_i) + g_o d_k = g_o d_k + x_k, \end{aligned} \tag{5}$$

since $h_i + g_i = 0,\ i \neq 0$, and since $b_i \neq d_i$ implies $b_i = d_i$, it follows that

$$x_k = \sum_{\substack{i=0 \\ b_i = -d_i}}^{k-1} (h_{k-i}b_i + g_{k-i}d_i) = 2 \sum_{\substack{i=0 \\ b_i = -d_i}}^{k-1} g_{k-i}d_i. \tag{6}$$

The cumulative effect of errors prior to time k ($d_i \neq b_i,\ i \leqq k-1$) is expressed by the real number x_k.

The equation of the system, namely (3) now becomes

$$b_k = \operatorname{sgn}\{n_k + x_k + g_o d_k\}. \tag{7}$$

Hence, the output b_k will be in error when

$$\begin{aligned} n_k + x_k &< -g_o \quad \text{if} \quad d_k = 1 \\ n_k + x_k &> +g_o \quad \text{if} \quad d_k = -1. \end{aligned} \tag{8}$$

Thus, the effective noise in the system at time k is not n_k but $n_k + x_k$.

Due to the assumptions concerning the independence of all input random variables d_k, the variables n_k and x_k are independent. Let the d.f. of x_k be $F_k(x)$. Also, let $p(k) = \text{prob}\,\{b_k \neq d_k\}$. It follows from (8) that

$$p(k) = p \int N(-g_o - x)dF_k(x) + q \int (1 - N(+g_o - x))dF_k(x). \tag{9}$$

Since p, q, and $N(x)$ are known, the problem reduces to the study of the sequence of r.v.'s $X_0, X_1, \cdots$ with d.f.'s $F_0(x), F_1(x), \cdots$.

It follows from (*iii*) and (6) that

$$\begin{aligned} x_{k+1} &= 2 \sum_{\substack{i=0 \\ b_i \neq d_i}}^{k-1} g_{k+1-i}d_i + 2g_1 d_k \\ &= rx_k - ad_k \quad \text{if} \quad d_k \neq b_k . \end{aligned} \tag{10}$$

where $a = -2g_1 > 0$. Or

$$x_{k+1} = rx_k \quad \text{if} \quad d_k = b_k . \tag{11}$$

Thus, there are three possibilities for transitions, each of which takes place with probability depending on the value of x_k. Namely,

$$\begin{aligned} x_{k+1} &= rx_k - a \quad \text{if} \quad d_k = 1 \neq b_k \quad \text{with probability} \quad p_1(x_k), \\ x_{k+1} &= rx_k \quad \text{if} \quad d_k = b_k \quad \text{with probability} \quad p_2(x_k), \\ x_{k+1} &= rx_k + a \quad \text{if} \quad d_k = -1 \neq b_k \quad \text{with probability} \quad p_3(x_k). \end{aligned} \tag{12}$$

The transition probabilities $p_n(x)$, $n = 1,2,3$ are determined from (8):

$$\begin{aligned} p_1(x) &= pN(-g_o - x) \\ p_2(x) &= 1 - p_1(x) - p_3(x) \\ p_3(x) &= q[1 - N(g_o - x)]. \end{aligned} \tag{13}$$

If we assume that the r.v. X_0 is independent of all the input variables, then the sequence of r.v.'s $X_0, X_1, X_2, \cdots$ where x_{k+1} is related to X_k by the transitions characterized by (12) and (13) forms a Markov chain. We shall use this property, however, only to the extent of relations (12) and (13).

Observe the way X_{k+1} is obtained from X_k. Given the value of X_k, the random variable X_{k+1} may have three possible values, the values assumed at X_k by three linear functions defined over the range of X_k.

The choice of the actual transformation used to generate the next value X_{k+1} of the r.v. X_{k+1} is made by performing an independent experiment with three possible outcomes with respective probabilities $p_1(X_k)$, $p_2(X_k)$ and $p_3(X_k)$ which, as indicated, are functions of the value X_k.

IV. ITERATIVE PROCEDURE FOR THE COMPUTATION OF ERROR PROBABILITIES

This type of random walk is studied in Ref. 1. It is shown there that whenever $0 \leqq r < 1$, the sequence of random variables $X_1, X_2, \cdots$ has a limiting distribution $A(x)$ and also that the mean value of any continuous function $f(x)$, with respect to $A(x)$, can be computed by iteration without actually obtaining $A(x)$. This result will now be applied to our problem.

Let $Uf(x)$ be the function

$$Uf(x) = p_1(x)f(rx - a) + p_2(x)f(rx) + p_3f(rx + a) \quad (14)$$

where $\quad p_n(x), \quad n = 1,2,3 \quad$ are given by (13).

Also, denote by $U^kf(x)$ the kth iterate of the transformation $Uf(x)$, namely,

$$U^kf(x) = U(U^{k-1}f(x)) \qquad k = 1, 2, \cdots . \quad (15)$$

Then, from Ref. 1 we have

$$\lim_k U^kf(x) = \int f(x)dA(x). \quad (16)$$

We then obtain from (9), (13), and (16) that

$$\lim_k p_k = \lim_k U^k(p_1(x) + p_3(x)). \quad (17)$$

For the general case of l consective errors starting with the kth output, we write

$$p(k,l) = \text{Prob}\ (b_k \neq d_k, \cdots, b_{k+l-1} \neq d_{k+l-1}) \quad (18)$$

and for the conditional probability of l consecutive errors given $X_k = x$, we write

$$p(k,l \mid x) = \text{Prob}\ (b_k \neq d_k, \cdots, \quad b_{k+l-1} \neq d_{k+l-1} \mid X_k = x). \quad (19)$$

Clearly,

$$p(k,l) = \int p(k,l \mid x)\ dF_k(x). \quad (20)$$

On the other hand, it follows from (12) and (13) that

$$p(k,l \mid x) = \sum_{\substack{\varepsilon_i \neq 0 \\ i=1,\cdots,l}} p(\varepsilon_1, \cdots, \varepsilon_l \mid x) \tag{21}$$

where

$$\begin{aligned} p(\varepsilon_1, \cdots, \varepsilon_l \mid x) = p_{2+\varepsilon_1}(x) p_{2+\varepsilon_2}(rx + a\varepsilon_1) \cdots \\ p_{2+\varepsilon_l}(r^{l-1}x + a(r^{l-2}\varepsilon_1 + \cdots + \varepsilon_{l-1})), \end{aligned} \tag{22}$$

with $\varepsilon_{i+1} = 1,0,-1$ according as $d_{k+i} = 1 \neq b_{k+i}$, $d_{k+i} = b_{k+i}$, $d_{k+i} = -1 \neq b_{k+i}$. Clearly, $p(k,l \mid x)$ is independent of k. Hence, on account of the theorem used before,

$$q(l) = \lim_k p(k,l) = \lim_k U^k p(k,l \mid x) = \int p(k,l \mid x) dA(x), \tag{23}$$

is the steady state probability of l consecutive errors. Other error patterns may be treated similarly.

V. SUMMARY

Observing that the cumulative error in a certain kind of data transmission system is a Markovian process we have derived an iterative procedure for computing the limiting probability of arbitrary error patterns. Using this method one can obtain numerical estimates of these probabilities by the aid of (23) once a computer program has been written to perform the iteration given in (14). Such a program is not presently available.

A more general treatment of data transmission systems in which error reduction is achieved by quantized feedback may be found in a paper by W. R. Bennett on Synthesis of Active Networks.[2]

VI. ACKNOWLEDGMENT

The author wishes to express his thanks to J. Salz for calling his attention to this problem.

REFERENCES

1. Zador, P. L., Random Walk in Compact Metric Space, to be published.
2. Bennett, W. R., Synthesis of Active Networks, Proc. Symposium on Modern Network Synthesis, New York, 1955.

Contributors to This Issue

G. W. Bleisch, B.S.E.E., 1951, University of Rhode Island; Bell Telephone Laboratories, 1951—. His early assignments included circuit development in the military systems area and systems engineering work concerned with Bell System data processing studies. Since 1961, Mr. Bleisch has headed groups at the Merrimack Valley Laboratory concerned with N3 carrier terminal circuit design and the system analysis of short and long-haul carrier terminals. Member, Tau Beta Pi, Phi Kappa Phi, IEEE.

L. A. Bowie, B.S.M.E., 1960, University of New Hampshire; M.S.M.E., 1962, Northeastern University; Bell Telephone Laboratories, 1960—. Mr. Bowie has been responsible for various equipment designs of short-haul carrier system. His current assignment includes the equipment design of an N-to-L Carrier Interface. Member, Tau Beta Pi, Phi Kappa Phi, A.S.M.E., Registered Professional Engineer in New Hampshire.

Martin G. Cohen, A.B., 1957, Columbia College; M.A., 1958, Ph.D., 1963, Harvard University; Bell Telephone Laboratories, 1964—. Mr. Cohen has been engaged in research in optical modulation techniques since joining Bell Laboratories. Member, American Physical Society, Phi Beta Kappa, Sigma Xi.

David J. Farber, M.E., 1956, M.S., 1960, Stevens Institute of Technology; Bell Telephone Laboratories, 1956—. Mr. Farber currently is engaged in research in computer systems and programming languages. He supervises the systems programming group in the Holmdel computer department. He has been secretary and executive board member of SHARE. Member, Sigma Xi, ACM.

Winfield J. Giguere, B.S.E.E., 1954, University of New Hampshire; Bell Telephone Laboratories, 1954—. Mr. Giguere has been involved in development work in VHF radio and high-speed pulse transmission. He is currently a supervisor in a development area responsible for electronic packaging of VF transmission and signaling equipment. Member, IEEE, Tau Beta Pi, Phi Kappa Phi.

EUGENE I. GORDON, B.S., 1952, City College of New York; Ph.D., 1957, Massachusetts Institute of Technology; Bell Telephone Laboratories, 1957—. Mr. Gordon has been engaged in research in plasma physics, microwave tubes, lasers, and light modulation. He is presently head of the optical device department. Member, American Physical Society, Phi Beta Kappa, Sigma Xi, IEEE.

RALPH E. GRISWOLD, B.S., 1952, M.S., 1960, and Ph.D., 1962, Stanford University; Bell Telephone Laboratories, 1962—. Mr. Griswold has been engaged in research in programming languages and the application of computers to non-numerical problems. Member, ACM, Phi Beta Kappa.

R. L. HANER, B.S.E.E., 1959, Tufts University; M.S.E.E., 1961, Northeastern University; Bell Telephone Laboratories, 1959—. Since joining Bell Laboratories, Mr. Haner has been engaged in the design of short-haul carrier terminal circuits. He currently supervises the N carrier terminal circuit design group at Merrimack Valley. Member, Tau Beta Pi, IEEE.

CARL W. IRBY, B.E.E., 1939, New York University; Bell Telephone Laboratories, 1929—. Mr. Irby was first concerned with the evaluation of mitigative measures applied to outside telephone plant to control and reduce degradation of the telephone plant or its performance by outside sources. During World War II his evaluation work was on radar systems and their appliques. Since 1947, Mr. Irby has been engaged in the engineering and evaluation of short-haul carrier systems, N, O, ON, N2, and N3. He is now participating in the field trial and evaluation of the new L4 carrier system. Member, Eta Kappa Nu.

IVAN P. POLONSKY, B.A., 1949, M.S., 1952, and Ph.D., 1957, New York University. Instructor at Queens College 1953–60; Bell Telephone Laboratories, 1960—. Since joining Bell Laboratories, Mr. Polonsky has been working on programming languages for electronic computers and applications of computers to non-numerical problems. Member, American Mathematical Society, Mathematical Association of America, Association for Computing Machinery, Phi Beta Kappa.

J. E. SAVAGE, S.B., S.M., 1962 and Ph.D., 1965, Massachusetts Institute of Technology; Bell Telephone Laboratories, 1965—. Mr. Savage has been concerned with problems in coding and communication theory. Member, Tau Beta Pi, Sigma Xi, Eta Kappa Nu, IEEE.

JAMES A. WATTERS, B.S.M.E., 1930, Montana State University; Bell Telephone Laboratories, 1930–32 and 1936—. Mr. Watters has worked on various open wire and cable carrier systems and on communication systems for the armed forces. His later work involved short-haul radio microwave systems, N2 and N3 carrier terminal equipment, and the N2 carrier repeatered line. Member, Tau Beta Pi, Scabbard and Blade.

I. E. WOOD, B.A., 1928 and Eng., 1929, Stanford University; Bell Telephone Laboratories, 1929—. Mr. Wood's early work was concerned with negative feedback amplifiers for the Transatlantic telephone cable. During World War II he worked on antiaircraft gun directors and on the Nike guided missile project. He is currently engaged in development of circuits for carrier terminal equipment in the exchange transmission laboratory. Member, Sigma Xi, Phi Beta Kappa, IEEE, Professional Group on Circuit Theory.

PAUL L. ZADOR, B.A., 1958, Balliol College, Oxford, England; Ph.D., 1964, Stanford University; Bell Telephone Laboratories, 1963—. Mr. Zador has been engaged in the study of various problems involving information theory.

B.S.T.J. BRIEFS

A $\frac{1}{4}$-Watt Si PνN X-Band IMPATT (IMPact Avalanche Transit Time) Diode

By T. MISAWA and L. P. MARINACCIO

(Manuscript received May 13, 1966)

An output of 250 mW CW at 12 GHz with an efficiency of 2.8 percent was obtained from an Si PνN diode.

Previously[1] we discussed characteristics of Si PνN IMPATT (IMPact Avalanche Transit Time) diodes. In that study the emphasis was primarily in understanding device properties rather than extending their output power. In those diodes, the input power had to be kept less than 2 watts to avoid device burnout. The average microwave power output was about 10 mW with the efficiency seldom exceeding 1 percent. However, by immersing the diode holder in liquid nitrogen, and thus increasing heat flow, it was shown that the diodes were capable of CW output on the order of 100 mW in the X-band (8.2 to 12.4 GHz). The main deterrent to increased output power was the fairly large thermal impedance from junction to package.

To relieve the heating limitations, we recently mounted similar diodes with junction side down[2] which allowed us to apply input powers as large as 10 watts without lowering the ambient temperature. Fig. 1 shows the output and efficiency vs input plot for the best diode. The oscillation was measured in the 50 mil high waveguide circuit used by De Loach and Johnston.[3] The breakdown voltage was 54 volts. At the highest input the current density is about 2,000 A/cm^2. The efficiency is still increasing. This is a feature of the pνn structure in contrast to the Read structure which shows efficiency saturation at lower current density.[4] The oscillation frequency at the highest input current was 12.04 GHz; as the input current was decreased the frequency decreased by 0.2 GHz.

The microwave output was observed on the Hewlett-Packard spectrum analyzer 851A/8551A. Fig. 2 shows the spectrum observed with another diode from the same slice. The width of the signal is less than the bandwidth of the spectrum analyzer (1 kHz).

It appears that a substantial further increase in efficiency and output power should be possible. The small-signal analysis of a theoretical model, which simulates the actual structure, indicates that the per-

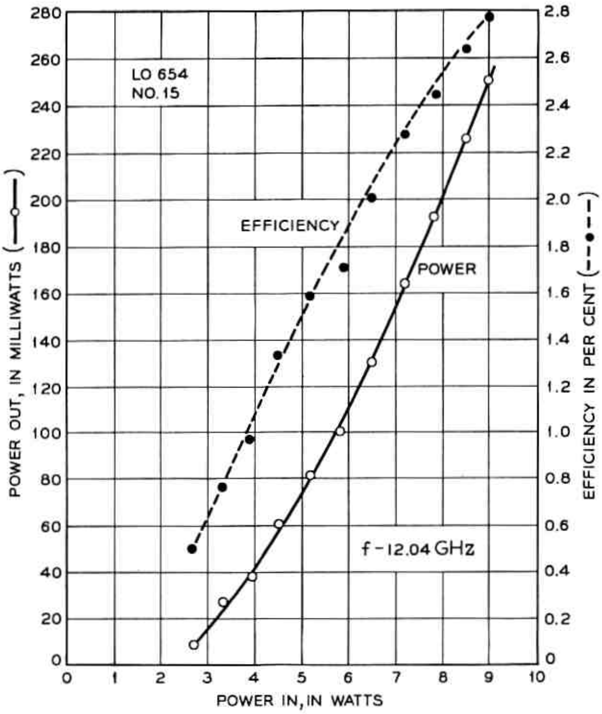

Fig. 1 — Output power and efficiency vs input power.

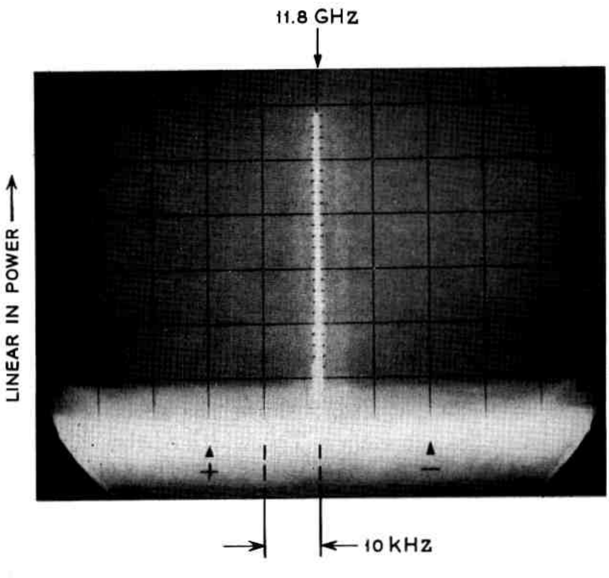

Fig. 2 — Spectrum of oscillation.

formance still keeps improving with increasing bias current. In addition, the thermal impedance calculated from the diode geometry and thermal conductivity of silicon should be much smaller than the observed value.

The authors would like to thank B. C. De Loach for many valuable suggestions and G. Pfieffer and A. Todd for fabricating the units.

REFERENCES

1. Misawa, T., Silicon Transit Time Avalanche Diode, IEEE Electron Devices Meeting, Washington, D.C., October, 1965 and to be published.
2. This method of mounting was used by De Loach and Johnston. Private communication.
3. De Loach, B. C. and Johnston, R. L., Avalanche Transit-Time Microwave Oscillators and Amplifiers, IEEE Trans. Electron Devices, *ED-13,* January, 1966, pp. 181–186.
4. Johnston, R. L. and Josenhans, J. G., Improved Performance of Microwave Read Diodes, Proc. IEEE (Correspondence), *54,* March, 1966, pp. 412–413.

Errata

A Note on a Type of Optimization Problem that Arises in Communication Theory, by I. W. Sandberg, B.S.T.J., *45*, May–June, 1966, pp. 761–764.

On page 763, replace the equation

$$\| u \| = \sum_{j \varepsilon \mathfrak{F}} | u_j |,$$

with

$$\| u \| = \sum_{j \notin \mathfrak{F}} | u_j |.$$

On the same page, replace equation (7)

$$\rho \triangleq \sum_{n \notin (\mathfrak{F} - \mathfrak{F}')} | \sum_{j \notin \mathfrak{F}} g_j x_{nj} | - \sum_{n \notin \mathfrak{F}} | \sum_{j \notin \mathfrak{F}} (g_j - c_j^*) \, x_{nj} | > 0 \qquad (7)$$

with

$$\rho \triangleq \sum_{n \varepsilon (\mathfrak{F} - \mathfrak{F}')} | \sum_{j \varepsilon \mathfrak{F}} g_j x_{nj} | - \sum_{n \notin \mathfrak{F}} | \sum_{j \varepsilon \mathfrak{F}} (g_j - c_j^*) x_{nj} | > 0 \qquad (7)$$