

THE BELL SYSTEM TECHNICAL JOURNAL

VOLUME XLII

MARCH 1963

NUMBER 2

Copyright 1963, American Telephone and Telegraph Company

The New L Multiplex — System Description and Design Objectives

By F. J. HALLENBECK and J. J. MAHONEY, Jr.

(Manuscript received October 24, 1962)

This paper discusses in broad terms the design of a radically new family of multiplex terminals, designated the L multiplex, designed to work with any of several broadband transmission facilities. The relevant historical background of multiplex terminals is covered and design objectives for the new equipment are outlined.

The new multiplex takes advantage of advances in the state of the art to reduce the size of equipment and increase reliability, while it retains the proven advantages of older equipment. Other advantages of the new multiplex include in-service maintenance, decentralized carrier supplies, and increased flexibility. The latter permits its use with as few as 60 or as many as 1860 voice channels and also adapts it for use with wideband data service. Compatibility with older multiplex equipment was maintained where possible, but in order to meet future demands for broadband service, it was decided to modify older equipment rather than compromise the design of the new to maintain compatibility.

I. INTRODUCTION

The coaxial cable and radio relay facilities of the Bell System employ the same form of frequency division multiplex. This paper describes the terminal multiplex arrangements in broad terms and discusses the design objectives for a radically new family of multiplex terminals having message channel capacities ranging from 60 to 1860. Dependent upon the A-type channel bank¹ for the first step of modulation from voice frequencies to group carrier spectrum and the reverse function of demodulation, the new multiplex accepts such groups and assembles them into a single broadband signal for transmission via the various high-

frequency media. It is an order of magnitude smaller than its predecessor and offers operating advantages not formerly available. The first commercial installations were placed in service in mid-1962. Circuit and equipment design features of the new multiplex are covered in detail in companion papers.^{2,3,4}

II. HISTORICAL BACKGROUND

2.1 *General*

Realization of the potential of a wire or radio system to transmit many messages simultaneously requires multiplexing equipment at the terminals of the transmission medium. Multiplexing techniques may be divided into two categories, frequency-division and time-division. Broadband facilities capable of handling hundreds of channels have used the former exclusively, and this paper is concerned with modern versions of such arrangements.

Essentially the transmitting circuits of a multiplex convert signals from a number of independent voice-frequency channels into a single broadband signal suitable for transmission over a high-frequency medium. The receiving circuits reverse this action by resolving the single broadband signal into individual VF channels. The circuit and equipment concepts for the earliest carrier systems were quite simple. With available lines severely limited in bandwidth, three or four channels were the maximum attainable. These were formed by a single modulation step using carriers which were provided by simple, individual oscillators.

As the art advanced, much broader bands of frequencies could be transmitted and the economics of system planning indicated that a basic new approach was needed. One outcome was the development of the A-type channel bank.

2.2 *A-Type Channel Bank*

In the early 1930's two carrier systems were under active development; one for 19-gauge toll cable pairs (Type K), one for open-wire pairs (Type J). Crosstalk coupling and line attenuation limited the frequency band so that only twelve channels could be transmitted. However, looming in the not too distant future was a system based on the new concept of a coaxial cable capable of transmitting a much wider band with channel potentialities 40 to 50 times greater.

To provide a common denominator for all of these systems, the A-type channel bank was developed. It became the first step of modulation and

yielded twelve single-sideband voice channels with suppressed carriers spaced at uniform 4-kc intervals in a standard group frequency band from 60 to 108 kc. This channel format is standard in the Bell System and has been adopted internationally. The historical background for this choice of channel format and the continuing development of channel banks from the A1 design to the new A5 transistor version are discussed in detail in a recent paper.¹

Next in multiplex progress was the development of equipment required to operate with channel banks to provide the hundreds of channels which coaxial and, later, radio systems could handle.

2.3 *Multiplex for Coaxial Systems*

In the Type K cable and Type J open-wire systems for which the channel bank was first utilized, the translation to line frequencies was a simple matter because the system capacity was limited to a single group. For the much higher capacity L1 coaxial system^{5,6} both single-step and two-step modulation were considered.

A two-step plan was adopted for the following reasons:

1. Selectivity requirements on the band filters were eased.
2. Fewer types of band filters were required.
3. Fewer carrier frequencies had to be produced.
4. As in the case of the standard group output of the channel bank, a large group of channels could be provided in a second common standard frequency range. These two provisions insured that flexibility was built into the multiplex to facilitate interconnection of systems without requiring reduction to voice frequency for all channels.

Study of traffic conditions and the economics of various arrangements led to the conclusion that in the second modulation step the output of five channel banks should be combined into a basic supergroup of 60 channels. This basic supergroup from 312 to 552 kc also became standard both in the Bell System and internationally. In the original L1 multiplex eight supergroups were combined for line allocations from 68 to 2044 kc. Later two supergroups were added at higher frequencies, which were intended only for shorter haul traffic due to transmission limitations of the line.

The multiplexing arrangements which are represented by this array of groups and supergroups and the necessary carrier supplies encompass the equipment involved in the new L multiplex as shown in Fig. 1.

About a decade after the L1 coaxial cable system became a reality, a

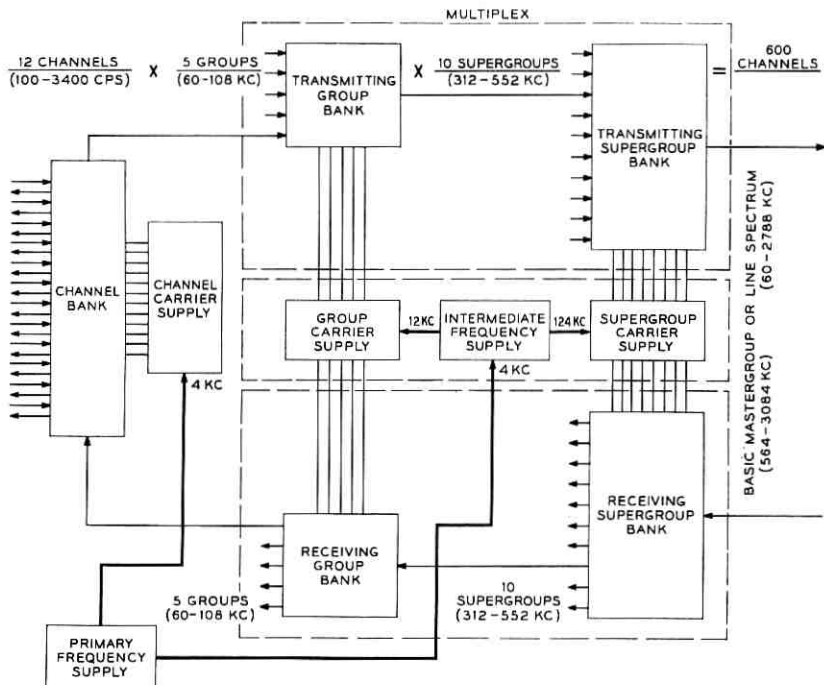


Fig. 1. — Broadband mastergroup; L multiplex equipment enclosed in dashed lines.

more complex multiplex was developed. This was required for the new L3 coaxial system⁷ which employed highly refined amplifiers and pilot regulators along with extremely accurate fixed and variable equalizers. With 4-mile repeater spacing instead of the eight miles of L1, the useful transmission band was extended to about 8 megacycles. A corresponding channel capacity of over 1800 channels was achieved. To attain this larger capacity a new concept of combining three mastergroups of 600 channels was introduced. The equipment developed for this final step of modulation has not yet been redesigned. However, the new L Multiplex does provide the necessary supergroups to form a basic mastergroup for the L3 terminals.

2.4 Multiplex for Microwave Radio

The development of multiplex for wire systems antedated microwave radio systems by many years. However, with the design of the first commercial long-haul microwave radio system, TD-2,^{8,9} it was soon evident that the earlier multiplex developed for wire systems would be

satisfactory for radio terminals. The 600 single-sideband, suppressed carrier channels of the L1 multiplex matched the load capacity of a TD-2 broadband channel for many applications. The use of the same multiplex for radio and wire systems offered significant benefits in standardization. For example, it permitted efficient and flexible interchange of traffic at offices using both types of facilities.

As microwave developments progressed, the standard terminal pattern was followed. The latest long-haul radio system, TH,¹⁰ employs the L3 mastergroup multiplex (1860 channels) on each broadband radio channel. The new short-haul lighter route radio systems, TJ¹¹ and TL, use partial L1 multiplex arrangements up to their maximum load handling capabilities.

2.5 Carrier Supply

The decision to adopt uniformly spaced channels based on harmonics of 4 kc¹ set the pattern for a common carrier supply. The earlier small-capacity systems for use on paired cables and open-wire lines used a primary generator based on a tuning fork to supply the 4 kc. Harmonics of 4 kc were formed by driving a saturable reactor as a pulse generator. The needed carriers and pilots were selected by filters. When the L1 multiplex was developed, many additional carriers and pilots were needed. Although the same general method was followed, certain alterations were made in the plan.

To obtain the higher absolute accuracy needed, the 4-kc base frequency was derived as a subharmonic from a high-stability crystal oscillator operating at the favorable crystal frequency of 128 kc. As before, the channel and group carriers were obtained from a 4-kc harmonic generator. For the supergroup carriers, however, a new 124-kc harmonic generator was added with this drive frequency derived from the 4 kc base. Frequency differences among terminals working together to provide certain special services such as VF telegraph and program were maintained to less than about one part in 10⁷. Frequency precision of this order limits the shift in all channels to less than 2 cps. This was accomplished by a system of master and slave 128-kc oscillators controlled by a standard reference frequency originating at New York.

III. STANDARD ARRANGEMENTS — NEW L MULTIPLEX

3.1 Nomenclature

A large-capacity multiplex is assembled from many repetitive units which may be considered building blocks. Each complete multiplex is an

assemblage of such blocks uniquely suitable to terminate a particular broadband facility. With the introduction of a new multiplex family, it is desirable to use a readily understood descriptive designation for each complete multiplex. The new multiplex design is radically different, but the basic system plan has been retained and the multiplex applications are identical to those of the older equipment. In view of these latter factors and the field familiarity in referring to this equipment as the "L" carrier terminal, it was decided to retain "L" as part of the various general designations. In this usage the letter L implies single-sideband channels spaced at 4-kc intervals and assembled in the standard group, supergroup, and mastergroup format.

3.2 *Standard L Multiplex Combinations*

In the older L carrier terminal, designations such as "group bank," "supergroup bank," and "carrier supply" were used. These were functionally descriptive and were logical separations since assemblages of these functional units occupied many bays of equipment. With the very great size reduction achieved in the new multiplex, these lines of separation become blurred. For example, only one transmit and one receive bay are needed to provide the group, supergroup, and carrier supply equipment for a complete 600-channel terminal. These considerations led to the descriptive coding that has been applied to the standard L multiplex family. (J coding is applied as usual for the various sub-units for manufacturing and ordering purposes.)

The various codes are:

L600A — This is the multiplex to convert as many as 50 standard groups to line frequency allocations for L1 coaxial and TD-2, TJ, and TL radio. A maximum of 600 channels is available, but smaller numbers can be provided.

L1860A — This is the multiplex to convert as many as 155 standard groups into the basic mastergroups of the L3 coaxial and TH radio systems. A total of 1860 channels (3 mastergroups + 1 supergroup) are available but fewer supergroups can be furnished. These bays do not include the final step of modulation for translation from basic mastergroup to line frequency allocations.

L60A and L120A — These differ in principle from the equipments listed above in that they are complete packages providing channel banks and voice patching units as well as the group, supergroup, and carrier supply. A total of 60 or 120 channels is available but smaller numbers can be provided. These small packages are intended to provide economical terminals for light route TJ or TL

radio systems or in other special instances where ultimate traffic needs are not expected to approach the capacity of an L600A.

IV. PERFORMANCE OBJECTIVES

4.1 *General*

The ultimate objective for performance is that the new multiplex be suitable for the service demands of today and for those of the foreseeable future. One of today's more stringent demands has been brought about by customer Direct Distance Dialing. In manual operation, the end product of transmission service offered to the telephone subscriber had been inspected by the operator. In other words, a long distance telephone connection was offered to the subscriber only after the operators had conducted a satisfactory conversation over it. In Direct Distance Dialing, the subscriber himself is the first person who attempts conversation over the particular connection set up for his use by machine operation. The lack of inspection of the end product then requires much stricter maintenance of the individual facilities for adequate assurance that they will provide satisfactory service each time they are assembled in tandem to complete particular telephone connections. The number of subunits likely to be connected in tandem is increased by the complex plant layouts which make efficient use of the broadband facilities. A further increase occurs from automatic alternate routing during heavy traffic periods.

Transmission performance of the original multiplex has been generally satisfactory, and the broad objective for a new design is that its comparable performance be at least equally good. In view of the increasing emphasis on multilink operation, however, improvement in certain critical characteristics such as terminal noise and deviations from uniform frequency response across groups and supergroups would be desirable to the extent that it is economically feasible.

4.2 *Maintenance*

One important result of maintenance studies conducted over the past several years has been to demonstrate that the basic design of multiplex equipment of the same type now in the telephone plant is sound. The equipment is inherently stable in most respects, but the actual control of the net loss of telephone channels is dependent on skilled personnel to carry out a coordinated program of elaborate test and adjustment procedures. Adjustment of the flat transmission of units handling only a

narrow band of frequencies is provided to compensate for departures from ideal flat transmission through units in which several of the narrow bands are handled as a single broadband of frequencies. For example, individual supergroups are adjusted to be equal at only one frequency to compensate for the frequency characteristic of the high-frequency line or radio baseband transmission. Similarly, individual groups are adjusted to be equal at only one frequency to compensate for the frequency characteristic across the supergroup. Finally, individual channels are adjusted to be equal at only one frequency to compensate for the frequency characteristic across the group band. In this way, frequency characteristics of the broader bands are equalized segment by segment. A particular set of adjustments, however, is appropriate only for the assemblage of equipment in operation at the time the final adjustments were made. Patching service to an alternate facility or substitution of spare equipment units will disturb the equalization. In fact, such mobility in the plant was found to be a principal cause of net loss deviations of individual message channels.

It also became apparent that with the increasing complexity of the telephone plant, manual maintenance methods were becoming inadequate. Accurate control of terminal transmission was dependent on the fine-grain adjustment of gains and losses of many subunits connected in tandem, where the adjustment of any one interacts with the adjustments of the others. A major difficulty was found to be coordination of the test and adjustment activities of maintenance people who may be in different offices separated by hundreds of miles. These problems point up the need for equipment design for which maintenance is less dependent on the skill of the craftsman and less demanding of his time.

Studies and feasibility trials in which new features were actually appended to the original multiplex had also demonstrated that the maintainability of a multiplex system would be enhanced by the use of terminal regulation to automate routine measurement and adjustment, and by simplifying those tests which must continue on a manual basis. By making adjustments the instant they are needed, automatic regulation minimizes the need for coordination of other adjustments. A man making a manual adjustment can feel confident that a preceding level is always being held at proper value. In other words, the usefulness of the adjustment he has just made is not apt to be destroyed by the delayed reaction of a craftsman elsewhere in the system. Regulation also provides the obvious advantage of freeing skilled manpower for maintenance tasks more demanding and less predictable than routine measurement and adjustment of pilot levels.

4.3 *Service Requirements*

Requirements for telephone service are growing more severe in at least two different ways. First, because telephone circuits are automatically made up by switches from a random assemblage of many units, individual units must be held to very close tolerances to insure a well controlled over-all net loss. Thus, requirements for precision of measurement and adjustment have become more severe than had been acceptable in the past. To meet these requirements, specific purpose test equipment is needed as an integral part of the multiplex to facilitate routine measurements of pilots in which simple and fixed selectivity, in-service test access through interlocked switching and readout on expanded scale meters are desirable design objectives. It is important also to provide convenient in-service adjustments located so that the meter can be read while making the adjustment. Meter readings should be in terms of departure from nominal to obviate the need for remembering many unique combinations of frequencies and test levels. Finally, the use of terminal regulation requires warning alarms when the necessary terminal adjustment becomes excessive. This warning will generally call for action at a location other than the receiving terminal because indications are that the levels arriving there are so far from nominal that further terminal correction will probably result in impaired performance. A pilot failure alarm is also a necessary feature to call prompt attention to a group or supergroup whose transmission has been disrupted. This feature can be had by moderate addition to the basic circuitry necessary for continuous regulation that is responsive to change in pilot output power.

A second service requirement that is growing more severe is the need for continuity of service, which will be intensified by new service offerings that encourage the use of the telephone plant outside of regular business hours to transmit data. Thus, modern equipment should not require disturbance of service merely to measure transmission accurately and to make any necessary adjustment. Therefore, a new multiplex should provide access for in-service measurement and adjustment. Hybrid-derived dual outputs have been provided in receiving units of the original multiplex to facilitate routine, in-service measurement of pilots to guide terminal adjustments. The independent hybrid output is essential to accurate power measurement to avoid the risk of serious error when power is deduced from bridging voltage measurements across impedances that are not accurately known. Similar test access is also necessary in transmitting units to facilitate maintenance of like units when it is necessary to substitute a spare for a regular unit with minimum disturbance of service.

4.4 *New Services*

In addition to providing for message service and the many forms of high-speed data services that can be handled by message channels, there is a growing need for facilities to handle even higher speed data requiring bandwidths much wider than a single message channel. Thus, a new multiplex must include terminal connectors to accommodate services which demand the wider bands available at the standard group and supergroup access to the multiplex. Group data services will displace 12 message channels; supergroup data service will displace 60 message channels.

V. PHILOSOPHY OF DESIGN

5.1 *General*

The original multiplex design dates back more than twenty years, and it was quite obvious that substantial benefit could be realized from nothing more than redesign of equipment to exploit modern components and design techniques. That simple approach, however, would neglect the importance of maintenance and fail to recognize new requirements. The actual philosophy of design has been much broader with the main purpose of producing a family of multiplex arrangements to efficiently meet the needs of today and those of the foreseeable future. The following points have been emphasized:

1. Exploit advances in the art.
2. Retain features of proven integrity in the design now in use.
3. Correct known deficiencies of the design now in use.
4. Where design differences would threaten compatibility of new equipment with old, reasonable modification of equipment in plant should be undertaken in preference to accepting compromise in the new design.

An outstanding example of advance in the art is the use of new ferrite materials in components to effect dramatic reduction in the size of filters. This is especially significant since filters represent an important part of the total bulk and cost of a frequency-division multiplex.

Design features whose operational value has been proven by years of experience include the use of single-sideband channels with suppressed carrier, standard frequency allocations for basic groups and supergroups, and the use of pilots to monitor group and supergroup transmission. All of these worthwhile features should be retained in a new design.

The multiplex equipment now in use is known to be deficient in ease

of maintenance to meet the strict demands of service today. As previously discussed, the new multiplex should include terminal regulation and in-service maintenance access to minimize this deficiency. Another fault is that the older equipment was designed for heavy cross-section use and has not been economical where comparatively few channels are needed. A new multiplex design should include decentralized carrier supply, as discussed in the next section, to overcome this restriction in application.

5.2 Carrier Supply

In the original multiplex design, carrier power to drive the numerous modulators used in frequency translation was derived from a centralized block of interrelated equipments. Such a centralized carrier supply is economical when fully loaded because its cost is shared by a large number of telephone channels. However, this high cost of common equipment is a serious deterrent at small terminal applications. In a new design, it is practical to consider generating carrier power in small equipment units to be mounted close to the point of use. Such a decentralized carrier supply minimizes the cost of getting started and enables the addition of carrier supply capacity in smaller increments as needed. A supplementary benefit of decentralization is to minimize the bulk of interbay cabling. This is possible because it is sufficient to distribute a single base frequency in place of the several individual harmonic frequencies needed for carriers in modulation. The objective is to decentralize carrier supply to the extent that small cross-section terminals become more economical, without significant cost penalty at the heavy cross-section terminals.

5.3 Compatibility

It is essential that multiplex equipment of new design be compatible with that of older design because, in the normal process of plant extension, it will sometimes be necessary for a new terminal to work with one of older design at the opposite end of the high-frequency transmission medium. Even when new and old designs of terminal are used in juxtaposition in the same office, they must be operationally alike to facilitate maintenance and enable service restoration in emergencies. At the expected high rates of future production, however, the penalty of compromise in the new design to retain compatibility would apply to an indefinitely large number of units, whereas reasonable modification of older equipment to resolve incompatibility will apply to a definite number of units already in plant.

The rapid rate of growth of multiplex equipment in the Bell System plant is illustrated in Fig. 2. The curves show the number of broadband terminals shipped for L multiplex use over the past years. Broadband terminal is the unit used in programming factory production and includes proportionate shares of the multiplex equipments needed for one terminal having a message-circuit capacity of 12 channels. The lower plot shows shipments each year, and these fluctuate with business conditions. The upper curve is a plot of the cumulative total shipments, which shows a definite exponential trend in rate of growth that has persisted over the past ten years. The straight-line approximation corresponds to an annual rate of increase of about 20 per cent per year which implies that the total quantity in plant will double in about four years. It is this high rate of growth that has encouraged consideration of accepting the cost of modernizing the older equipment by modification to retain compatibility, rather than accept a compromise in the new design.

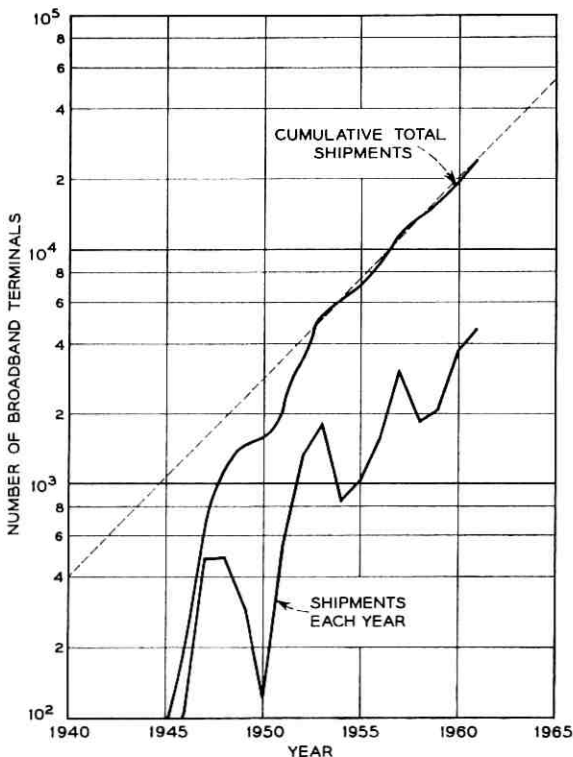


Fig. 2 — Broadband terminals shipped by Western Electric Co., 1945-1961.

An example of deliberate acceptance of incompatibility to be resolved by modification of the older plant is the new frequency allocation for Supergroup No. 1. The Bell System standard frequency allocation for Supergroup No. 1 has been the band from 68 to 308 kilocycles, and it is being changed to a new standard allocation from 60 to 300 kilocycles. This will result in additional guard space between the upper edge of Supergroup No. 1 and the lower edge of the basic supergroup, which eases substantially the design requirements on band filters for Supergroup No. 1. It is estimated that the long-term savings in filter cost for new terminals due to the change in frequency allocation will be greater than the total cost of modification of plant in service in a reasonably short time, and that substantial savings will accrue beyond that time. In addition to economy, there should also be an intangible advantage in minimizing conflict between frequency allocations standard in the Bell System and those recommended internationally.

A second example of incompatibility accepted for resolution by modification of plant in service is the use of a single transmitting cable from channel bank to group equipment. Hybrid-derived dual outputs have been provided at the transmitting carrier frequency side of A-type channel banks for test access and to facilitate in-service patching of group equipment. The hybrid coil has been an integral part of the channel bank with both outputs cabled from there to the high frequency patch bay. A single cable will be used by locating a miniature hybrid coil in the high frequency patch bays of existing installations and in the new transmitting multiplex bays. The use of a single transmitting carrier cable will result in significant reduction in cost of installation for each new channel bank, and these savings are expected to more than counteract the expense of modification where older type channel banks are to be used with the new multiplex. In addition to long-term economy, there will be considerable relief from office cable congestion, a matter of increasing concern and expense to the operating companies. The use of single cable also enables the design of a compact group distributing frame, a new equipment item that is needed urgently to reduce the time and expense of necessary plant rearrangement to accommodate growth and changing pattern of traffic.

5.4 Continuing Development

The companion papers previously cited describe the multiplex development completed to date. Studies are continuing to formulate design objectives for the development of other new equipment used in close conjunction with the multiplex. A matter of increasing concern is the need

to insure continuity of service under emergency conditions. Emphasis is being placed on new equipments for the following purposes:

1. To enable centralized control and maintenance of the carrier facilities.
2. To prevent the propagation of trouble into otherwise unimpaired facilities.
3. To accelerate prompt restoration of service.

Another important area of study is concerned with the operational integration of standard wideband data service with basic telephone message service. Study to date of the latter area has already led to the decision to change the standard frequency allocations of the terminal pilots of the L-type multiplex. At present, a 92-kc pilot is used to monitor the transmission of each group, and the Group 3 pilot translated to 424 kc is used to monitor each supergroup. These choices are satisfactory for message service in that pilots are located near the center of the band they monitor and correspond to zero frequency or 4000 cycles per second in adjacent message channels. Since message service does not require the transmission of frequencies close to either of these extremes, the pilots do not interfere with message service and allow in-service measurement and adjustment of group and supergroup with minimum reaction on service. When the wider bands are offered for data service, however, the presence of a pilot near the center of the band is a serious restriction. This presents a basic conflict of interest in that a pilot is considered essential to good maintenance and yet its presence near the center of the band is a deterrent in standard wideband data service. The conflict has been resolved by agreement to move the pilots close to the edges of the wider bands, using 104.08 kc for group pilot and the Group 1 pilot translated to 315.92 kc to monitor supergroup transmission. These choices appear to sacrifice 4 kc of available bandwidth but the actual waste is much less because delay distortion impairs the band edges so that they would not be of much value in data transmission. The offset of 80 cps from an exact multiple of 4 kc is planned to obviate the need for an expensive crystal filter to guard against interference from channel carrier leak. It is recognized that changing the entire Bell System plant to new standard-frequency allocations for terminal pilots will incur substantial expense to convert existing equipment. Nevertheless, integration of wideband and message service is considered essential to future efficient use of the plant, and present indications are that a major part of the cost of conversion will be recovered in filter savings resulting from the use of an offset frequency.

VI. CONCLUSION

The new family of L-type multiplexes will enable the operating telephone companies to meet efficiently the need for modern communications over the coaxial cable and radio plant of the Bell System. Significant reductions in space and power requirements along with regulation and other operational advantages are made available for slightly lower first cost of installed equipment. Long-term economy should result from improved service due to more effective maintenance with less expenditure of effort.

REFERENCES

1. Blecher, F. H., and Hallenbeck, F. J., *B.S.T.J.*, **41**, Jan., 1962, p. 321.
2. Graham, R. S., Adams, W. E., Powers, R. E., and Bies, F. R., *B.S.T.J.*, this issue, p. 223.
3. Albert, W. G., Evans, J. B., Jr., Ginty, J. J., and Harley, J. B., *B.S.T.J.*, this issue, p. 279.
4. Clark, O. P., Drazy, E. J., and Weller, D. C., *B.S.T.J.*, this issue, p. 319.
5. Abraham, L. G., *Trans. A.I.E.E.* **67**, 1948, p. 1520.
6. Crane, R. E., Dixon, J. T., and Huber, G. H., *Trans. A.I.E.E.* **66**, 1947, p. 1451.
7. Elmendorf, C. H., Ehrbar, R. D., Klie, R. H., and Grossman, A. J., *B.S.T.J.*, **32**, July, 1953, p. 781.
8. Grieser, T. J., and Peterson, A. C., *Elect. Engg.*, **70**, Sept., 1951, p. 810.
9. Roetken, A. A., Smith, K. D., and Friis, R. W., *B.S.T.J.*, **30**, Oct., 1951, p. 1041.
10. McDavitt, M. B., *Trans. A.I.E.E.*, **76**, Part 1, Jan., 1958, p. 715.
11. Gammie, J., and Hathaway, S. D., *B.S.T.J.*, **39**, July, 1960, p. 821.

New Group and Supergroup Terminals for L Multiplex

By R. S. GRAHAM, W. E. ADAMS, R. E. POWERS,
and F. R. BIES

(Manuscript received October 24, 1962)

L multiplex terminals provide the several stages of modulation required to translate telephone channels into the broad basebands which can be transmitted on long-haul systems. The group and supergroup equipment has been redesigned incorporating many new features. Pilot-controlled regulators are provided to improve performance. Plug-in modules and automated monitoring facilitate maintenance. The equipment is arranged to reduce office cabling and installation cost. By the use of transistors and other newly developed devices and components, a size reduction of 10 to 1 has been achieved and the required power greatly reduced.

I. INTRODUCTION

The L multiplex terminals are used for Bell System long-haul telephone circuits, including those transmitted over the L1 and L3 coaxial systems, the TD-2, TH, TJ and TL microwave systems. Single-sideband amplitude modulation is used for translating voice channels into the broad frequency bands which can be transmitted by these systems.

Several stages of modulation are required to form the complete baseband. As shown in Fig. 1, these stages are known respectively as the channel bank, group bank, supergroup bank, and mastergroup bank. This paper describes a new design of the group and supergroup equipment.

A companion paper¹ outlines the original development of telephone terminals for coaxial systems, their subsequent use on microwave radio systems, and the increasing need for better performance to accommodate new services such as direct distance dialing and data transmission. The development of channel banks including the new A5 channel bank is described in a recent paper.² The original development of the group and supergroup equipment was started more than 20 years ago.³ Since

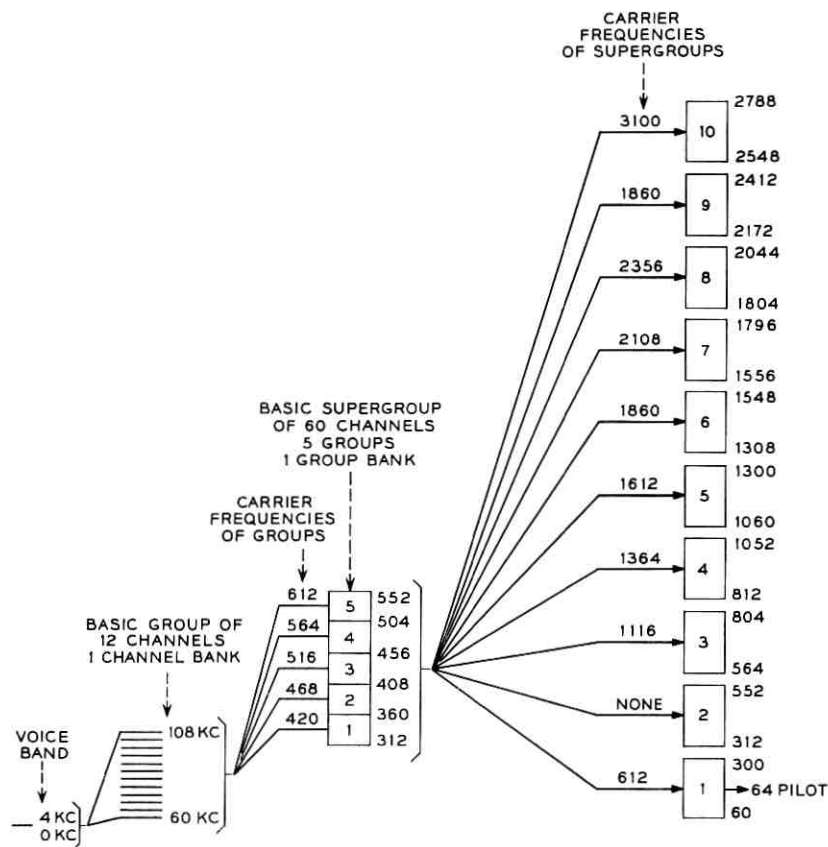


Fig. 1 — Frequency allocation — L600 multiplex.

then, the advent of the transistor and the development of associated components have made it possible to achieve a large reduction in size in the new design, even though several new features have been added.

1.1 *New Features*

Pilot-controlled regulators have been added to the receiving supergroup and group amplifiers. These improve over-all transmission stability by automatically correcting level changes which may occur anywhere along the circuit between the transmitting and receiving terminals.

All amplifiers employ transistors and are mounted in plug-in assemblies, introducing a high degree of flexibility and ease of installation and

maintenance. Automatic scanning and alarm circuits have been developed to facilitate maintenance and trouble location.

New group and supergroup carrier supplies for furnishing the required carrier frequencies for modulation, demodulation, and pilots have been developed. These are described in a companion paper.⁴

All equipment for translating the output of 50 channel banks (600 channels) into one mastergroup is arranged in one shop-wired transmitting bay. The corresponding receiving equipment is included in a second bay. This consolidation of equipment which formerly occupied more than 20 bays results in large savings in space and in interbay cabling. The power required has been greatly reduced, and only one supply, 24 volts dc, is required.

A group distribution frame for reassigning channel banks or group connectors has been developed to eliminate the need for recabling when changes are made in circuit layouts.

II. GENERAL DESCRIPTION

Two types of terminals have been developed, the L600 for multiplexing up to 600 voice channels, and the L1860 for multiplexing up to 1860 voice channels. The L600 is used as terminals for TD-2, TJ, and TL microwave radio relay systems and the L1 coaxial cable system. Its frequency allocation is shown in Fig. 1. The L1860 is used for L3 coaxial and TH radio systems. As indicated on Fig. 2, the L1860 employs different supergroup carrier frequencies for certain supergroups and requires the mastergroup stage of modulation. Equipment arrangements for smaller numbers of circuits are also being provided; these are the L60 and L120, for 60 and 120 channels, respectively.

2.1 *Frequency Allocation*

2.11 *L600 Multiplex*

The frequency allocations shown on Fig. 1 agree with those used on the L1 system telephone terminals except for that of Supergroup 1. The band shown, 60–300 kc, is 8 kc lower in frequency than that used in the L1 terminals. The new allocation permits the use of filters for Supergroups 1 and 3 which do not contain quartz crystals, which were previously required. The new Supergroup 1 carrier is the same frequency as the Group 5 carrier, 612 kc, permitting some simplification in carrier supplies. The new allocation agrees with the recommendation of the CCITT.*

* Comité Consultatif International Téléphonique et Télégraphique.

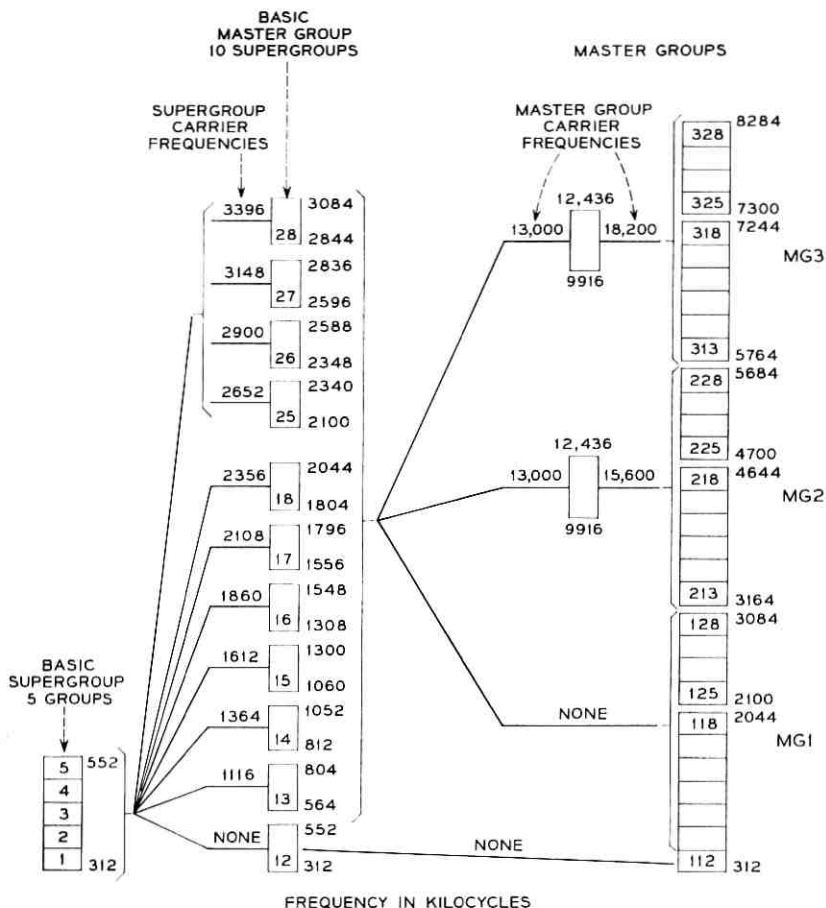


Fig. 2 — Frequency allocation — L1860 multiplex.

The continued use of 64 kc as a pilot frequency for microwave radio systems and for synchronization of carrier supplies makes it necessary to provide a 64-kc band-elimination filter in the transmitting Supergroup 1 to clear a band for the pilot.

2.12 L1860 Multiplex

The plan used to obtain the basic mastergroup for the L1860 multiplex, shown on Fig. 2, differs from that used in the L3 system telephone terminals.⁴ The four supergroups numbered 25 to 28 are modulated directly to their basic mastergroup allocations instead of using three

stages of modulation. The earlier plan required fewer supergroup carrier frequencies and fewer filter designs. However, with the use of the mastergroup terminals for both L3 coaxial system and TH radio system, it becomes economical to supply the additional carriers and filter designs in order to eliminate the sub-mastergroup stages of modulation.

2.2 *Transmitting Circuits*

The block schematic on Fig. 3 shows the group and supergroup circuits. The output of the transmitting portion of the channel bank is connected to terminals on the transmitting group distribution frame, and through the frame to the pilot insertion equipment on the group and supergroup transmitting bay. The 92-ke elimination filter clears a pilot channel between two voice channels. It is required to suppress the carrier leak at 92 ke which would interfere with pilot operation. The 92-ke group pilot* is introduced through a hybrid transformer. The output including the pilot is fed to a second hybrid transformer which provides two equal outputs. One of these is connected through normal contacts on jacks to the transmitting group equipment. The second output is available for monitoring or emergency patching.

If the terminal is retransmitting a group that has been received from an incoming system, the group signal is transmitted through a group connector including a bandpass filter, the group distribution frame and the pilot insertion equipment. If it is desired to operate with a through pilot, a pad is substituted for the pilot elimination filter and pilot hybrid transformer.

In small offices the group distribution frames may be omitted and the connections made directly from channel banks or group connectors to the group equipment.

A transmitting group amplifier is combined with the group modulator in a single plug-in assembly. The lower sideband of the modulator output is selected by a bandpass filter. The filters for the odd-numbered groups have their outputs in parallel. Similarly, the two even-numbered group filters are paralleled. The two sets of outputs are added in a hybrid transformer. The entire band from 312 to 552 ke (5 groups) is amplified in the transmitting intermediate amplifier. A low-pass filter to suppress out-of-band harmonic products follows, and two equal outputs are provided by a hybrid transformer.

Each group bank output is modulated by an appropriate carrier, and

* In order to accommodate wider bands for high-speed data transmission, development work has recently been started to change the group pilot to 104.08 ke and to use the new Group 1 pilot, 315.92 ke, for supergroup regulation.

one sideband is selected by a band-pass filter. An exception is Supergroup 2, which is not modulated, but transmitted directly. As with the group filters, the odd-numbered supergroup filters have their outputs paralleled on one side of a hybrid transformer and the even-numbered on the other side. One output of the hybrid transformer is normally terminated but can be used for monitoring. The second output is patched to a terminal trunk for connection to the transmitting system for the L600 multiplex. When two or three mastergroups are to be combined in an L1860 multiplex, the supergroup bank output is amplified by a transistor feedback amplifier with a gain of 26 db.

2.3 *Receiving Circuits*

At receiving terminals the baseband signal is received from the radio terminals or the mastergroup equipment. It is separated into supergroup bands by filters, and all except Supergroup 2 are demodulated to the basic supergroup band, 312–552 kc. Each supergroup signal is amplified by a regulated amplifier similar to that of the receiving group except that the Group 3 pilot at 424 kc is used to control the regulator.

The receiving group equipment accepts the 312–552-kc band from the supergroup equipment and separates it into group bands by filtering. The output of each group filter is demodulated to the basic group band, 60–108 kc. The lower sideband is selected by a low-pass filter and amplified. The gain of the amplifier is controlled by a regulator operated by the 92-kc pilot. This maintains the pilot at the output of the amplifier at a substantially constant level. Since the gain of the amplifier is constant over the group band, the regulation compensates for any flat level deviation that may occur in the transmission circuit, at either terminal, or in the connecting system.

The main output of the group amplifier is connected through normal contacts on jacks to the receiving group distribution frame and through it to the receiving portion of a channel bank or to a group connector. A second output of the group amplifier is connected to a scanner circuit and a pilot alarm circuit. These circuits will be described in later sections.

III. REGULATORS

In the L1 and L3 terminals, group pilots at 92 kc are added at the transmitting terminals. These are used for circuit monitoring and for periodic manual adjustment of receiving group gain controls. The pilot in Group 3, 424 kc, is used for adjusting the supergroup gain control. With the increased complexity of the toll plant, in which some group

circuits may be connected through several different transmission systems and several sets of terminals via group or supergroup connectors before reaching the receiving channel bank, small deviations in level frequently accumulate. With the more stringent requirements of Direct Distance Dialing and data transmission, it has become difficult to achieve satisfactory performance with manual adjustments.

Continuous adjustment of receiving supergroup and group gain by pilot-controlled regulators offers substantial improvement. After considerable field experience with experimental and prototype regulators installed in L1 terminals, it was decided to include both group and supergroup regulators in the new terminals.

3.1 Design Considerations

As indicated in Fig. 4, the regulator μ circuit (group and supergroup) consists of a thermistor, transmission amplifier, and portions of a dc amplifier. The regulator β circuit comprises a pilot bandpass filter, a control amplifier, and the remaining portions of the dc amplifier, including the reference. In accordance with fundamental feedback theory,⁶ the response of the regulator to variations in input signal is simply:

$$R(\omega) = \frac{1}{1 - \mu\beta} \quad (1)$$

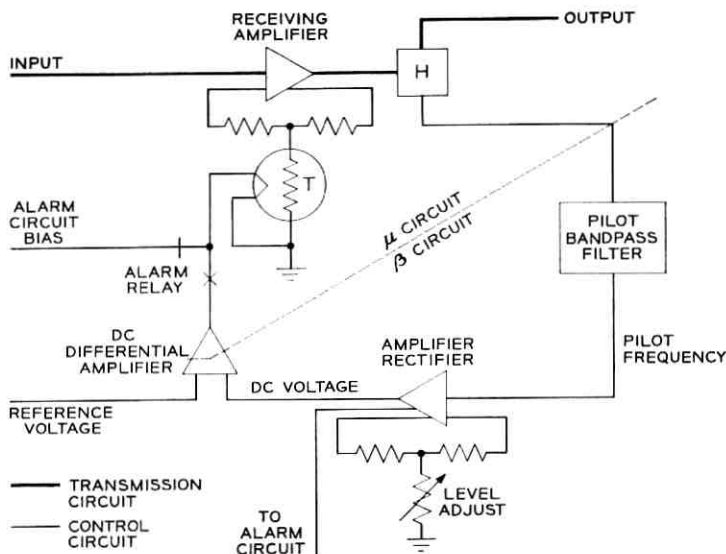


Fig. 4 — Block schematic of pilot-regulated amplifier (group or supergroup).

The value of (1) at dc envelope frequency is the compression of the regulator. A regulation compression of at least twenty to one was a basic design objective. As indicated in (1), such compression requires a comparable β -circuit expansion, or envelope gain. The expansion is provided by the dc amplifier operated against a stabilized diode reference.

System studies indicated that as many as ten regulators in tandem may occur on long routes with the greater proportion being supergroup regulators. In tandem operation, requirements placed upon the individual regulators are quite severe.^{7,8,9} Stability is readily analyzed in terms of the envelope frequencies associated with the controlling pilot.

In setting regulator design objectives, the total system envelope gain enhancement is a determining factor. Experience in regulator and servo-control system design had indicated that an over-all envelope gain of 3 db is reasonable for satisfactory transient and frequency response. Thus, for 10 units in tandem, the envelope magnification of each individual regulator should be limited to about 0.3 db to satisfy over-all performance.

A simple and direct method of minimizing gain enhancement is to cut off the $\mu\beta$ loop at 6 db per octave.⁷ The 6-db cutoff, to be effective, must be maintained through several octaves of envelope frequency. Thus, the loop gain should be well below zero before other circuit parameters force the cut-off to a more rapid rate. The circuit parameters required to minimize gain enhancement can readily be determined with the aid of a mathematical model.

As indicated in Fig. 5, the model consists of (i) a gain controlling element with a low-pass filter characteristic, (ii) an amplifier to provide envelope loop gain, and (iii) an equivalent envelope low-pass filter simulating the pilot bandpass filter.

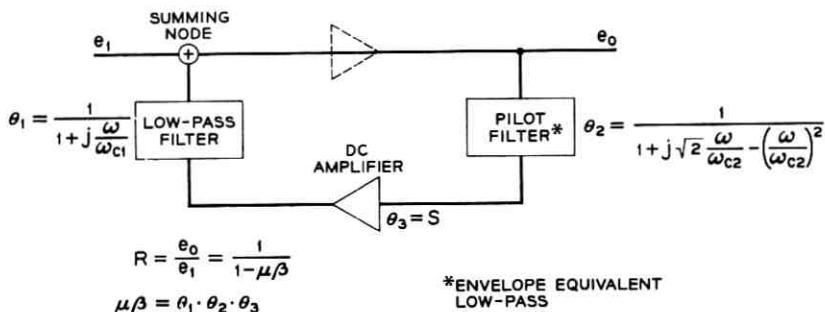


Fig. 5 — Mathematical model of regulator for computing envelope response.

The controlling element is represented as a simple lowpass filter with 6-db per octave attenuation characteristics. The low-pass envelope equivalent of the pilot bandpass filter may be represented as having 12-db per octave attenuation characteristics. The dc amplifier provides a loop gain, the cutoff characteristics of which may be neglected in an envelope analysis. From Fig. 5, it follows that the envelope loop gain is simply:

$$\mu\beta = \frac{S}{\left(1 + j \frac{\omega}{\omega_{c1}}\right) \left[1 + \sqrt{2}j \frac{\omega}{\omega_{c2}} - \left(\frac{\omega}{\omega_{c2}}\right)^2\right]} \quad (2)$$

where S is the envelope loop gain or stiffness, ω_{c1} the (radian) cutoff frequency of the low-pass filter, and ω_{c2} the cutoff frequency of the bandpass filter equivalent.

The regulator closed-loop frequency response follows from (1) and (2):

$$R(\omega) = \frac{1}{1 - \frac{S}{\left(1 + j \frac{\omega}{\omega_{c1}}\right) \left[1 + \sqrt{2}j \frac{\omega}{\omega_{c2}} - \left(\frac{\omega}{\omega_{c2}}\right)^2\right]}} \quad (3)$$

Analyses were made of (3) over a wide range of S and of cutoff frequencies. A convenience in such analyses is the parameter C , defined as the ratio of ω_{c2} to ω_{c1} . The results indicated an approximate relationship for the maximum value of $|R|$ in db, denoted by G , and the parameters S and C . For design, it is useful to consider C the dependent variable:

$$C \cong 11.6 \frac{|S|}{G} \quad (4)$$

The range of values for which (4) is valid is indicated in Fig. 6, which provides a useful regulator design chart.

A loop gain of 35, providing a compression of $\frac{1}{36}$, is a conservative limit for expansion. A desirable limit for gain enhancement is 0.3 db. Thus, referring to Fig. 6, the resulting value of C is 1350.

The computed response of such a regulator over a wide range of envelope frequencies is shown in Figs. 7 and 8. The loop gain ($\mu\beta$) cuts off at 20 db per decade (6 db per octave) until it approaches the cutoff frequency of the pilot filter. Loop phase shift first asymptotically approaches 90° , then later is further shifted by the phase contribution of the pilot filter. The response, R , is a 36 to 1 compression at low frequency, later falling off to 1, a decade before the pilot filter cutoff. At about 0.3 to 0.4 of the pilot filter cutoff, the response experiences a peak of ex-

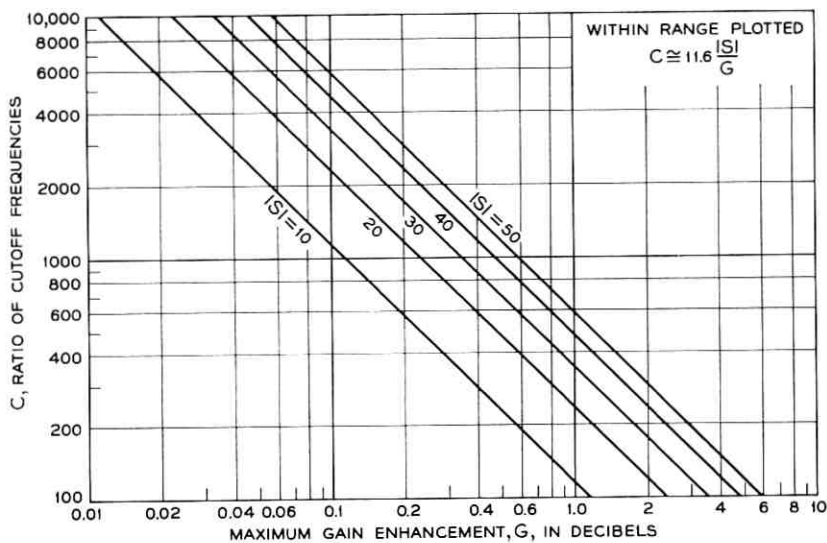


Fig. 6 — Relations among regulator design parameters.

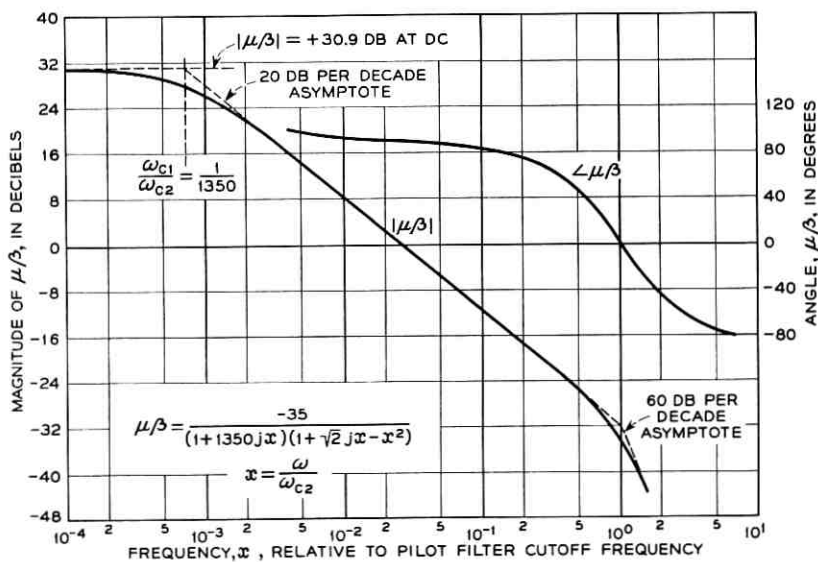


Fig. 7 — Computed loop-gain response of regulator.

pansion, which is shown in Figure 8 in exaggerated scale. As predicted by (4), the peak is about 0.3 db.

The time response of the model is readily determined by analyzing the appropriate Laplace transform, using analog computer techniques. Simplifying (3) and adopting the transform operator p for normalized frequency, the transfer function for the envelope response of each regulator becomes

$$R = \frac{Mp^3 + Np^2 + Qp + 1}{Mp^3 + Np^2 + Qp + (1 - S)} \quad (5)$$

where

$$\begin{aligned} M &= C & p &= j\omega/\omega_{c2} & N &= \sqrt{2} C + 1 \\ Q &= C + \sqrt{2}. \end{aligned}$$

The time response of the system to a unit step, $1/p$, at the input of the first regulator is given by the inverse transform

$$R_t = \mathcal{L}^{-1} \left[\frac{1}{p} \left(\frac{Mp^3 + Np^2 + Qp + 1}{Mp^3 + Np^2 + Qp + (1 - S)} \right)^n \right] \quad (6)$$

where n is the number of units in tandem.

Since the operator p was defined in terms of frequency relative to ω_{c2} , the time scale of R_t is restored by dividing (6) by ω_{c2} .

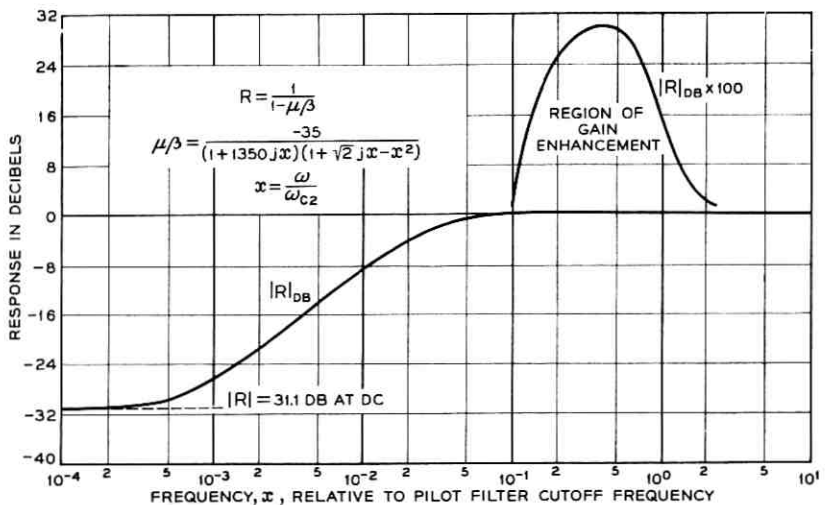


Fig. 8 — Computed through-transmission response of regulator.

Equation (6) is most readily solved using analogue computer techniques. A computed solution for the initial design parameters ($S = -35$, $C = 1350$) is shown in Fig. 9. It will be noted that, for 9 units in tandem, the overshoot is about one-half the amplitude of the initial step.

In evaluating time response, it is convenient to have as a reference the time response of a system with zero gain enhancement. Such a standard may be developed by assuming an infinite-bandwidth filter, i.e., only one source of energy storage in the control loop

$$\mu\beta = \frac{S}{1 + j\frac{\omega}{\omega_{c1}}} = \frac{S}{Tp + 1}$$

$$R = \frac{1}{1 - \mu\beta} = \frac{\frac{T}{1 - S}p}{\frac{T}{1 - S}p + 1} + \frac{1}{1 - S} \left(\frac{1}{\frac{T}{1 - S}p + 1} \right). \quad (7)$$

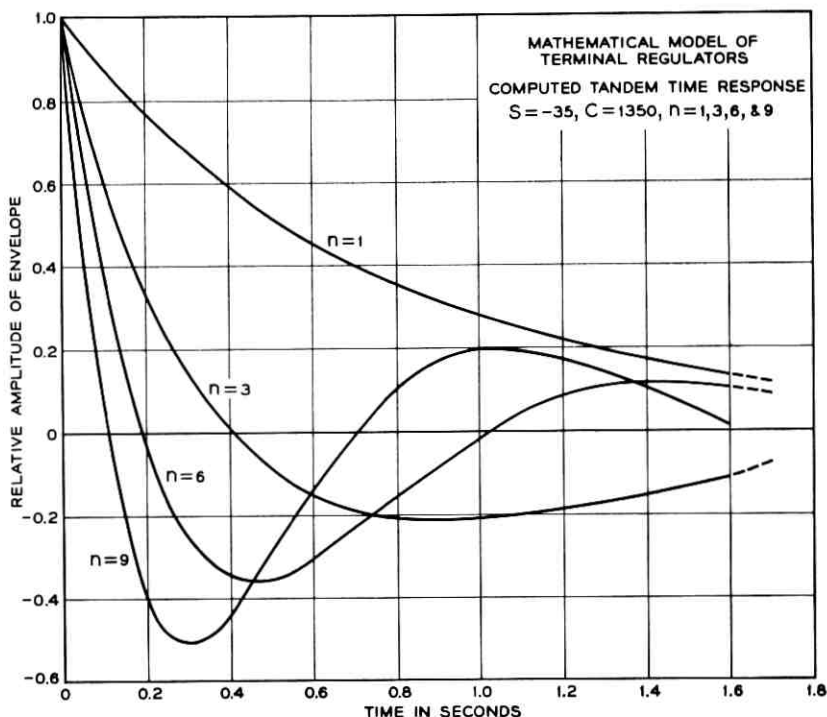


Fig. 9 — Time response of tandem regulators, computed.

If the $1 - S > 20$, the second part of (7) is negligible within the pass-band. In practice $1 - S$ is thirty or more; hence the response of the completely damped system is essentially that of a simple high-pass filter with 20-db per decade slope

$$R \cong \frac{p}{p + \alpha} \quad (8)$$

where $\alpha = (1 - S)/T$.

The time response of a number, n , of such systems in tandem is simply the inverse transform of

$$R_t = \mathcal{L}^{-1} \frac{1}{p} \left(\frac{p}{p + \alpha} \right)^n. \quad (9)$$

The general solution of (9) may be shown to be

$$R_{t(n)} = \epsilon^{-\alpha t} \sum_{k=0}^{n-1} \frac{C_{n-1,k} (-\alpha t)^k}{k!} \quad (10)$$

where

$$C_{n-1,k} = \frac{(n-1)!}{(n-1-k)! k!}.$$

Equation (10) will be recognized as a modified binomial expansion of $(1 - \alpha t)^{n-1}$, a fact that simplifies its numerical evaluation. The time response of completely damped systems was readily obtained by applying digital computer techniques to (10). Relative amplitudes of maximum overshoot — between the completely damped system and the model whose response is shown in Fig. 9 — are compared in Fig. 10. It will be noted that, for 9 units in tandem, the overshoot of the completely damped system is about 33 per cent, whereas that of the model is about 50 per cent.

3.2 Regulator Features

The mathematical model described above provided the basis for designing supergroup and group regulators for the L multiplex receiving terminal. However, the group and supergroup regulators differ in pilot frequency, bandwidth, and transmission levels.

The group regulator operates in the group baseband of 60–108 kc; its controlling pilot is 92 kc. The supergroup unit transmits the supergroup baseband, 312 to 552 kc; the pilot is 424 kc, the translation of 92 kc into Group 3. The output transmission level for group circuits is -5 db; for supergroup, -28 db, referred to zero transmission level.

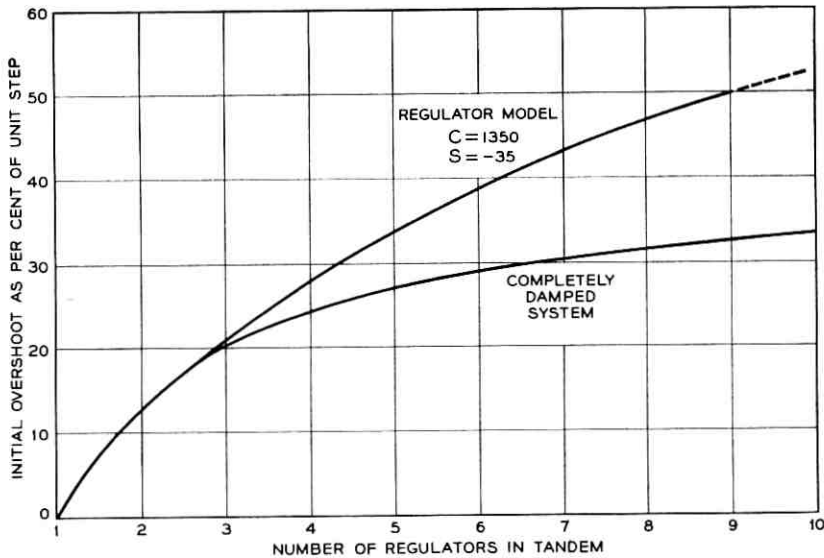


Fig. 10— Computed overshoot of terminal regulator compared with that of a completely damped regulator.

An indirectly heated thermistor (type 2A) was selected as the element controlling regulator gain in both units. The thermistor, with a time response of 60 seconds, approaches the type of low-frequency cutoff indicated as necessary by the mathematical model.

The thermistor is in the regulator μ circuit, and variations in bead resistance with temperature are regulated out by the control circuit. However, these variations do change the operating range of the thermistor and must be allowed for in control circuit designs. Also, the thermistor operating range affects the over-all expansion or envelope loop gain.

Most of the thermistor thermal inertia is concentrated in the heater. However, the bead does provide a separate, if smaller, source of energy storage. Thus at higher envelope frequencies the thermistor attenuation characteristic exceeds 6 db per octave and phase shift exceeds 90° . Envelope phase shift in excess of 90° contributes to gain enhancement.

In both group and supergroup regulators, the thermistor controls transmission gain by shunting the major loop feedback in the transmission amplifiers. Thus, the control method is the same for both units, i.e., amplifier gain varies directly as thermistor control current. The transmission amplifiers do not contribute appreciable phase shift within the pertinent band of the pilot envelope frequencies.

The proximity of adjacent transmission signals requires that the group and supergroup pilot bandpass filters provide sharp attenuation on both sides of the controlling pilot. The attenuation of both filters exceeds 40 db, two hundred cycles from pilot. The cutoff frequency of the group filter is about 15 cycles from pilot while that of the supergroup filter is approximately 50 cycles. The supergroup filter thus injects considerably less phase shift into the regulator loop within the pertinent band of envelope frequencies.

Amplification of the pilot within group and supergroup control loops is provided by a three-stage amplifier, tuned to the pilot frequency. The tuning is sufficiently broad that envelope phase shift is not appreciably affected. Output of the amplifier is rectified to provide about 4 volts dc across the 6000-ohm input impedance of the dc amplifier. A voltage doubler circuit on the output of the supergroup tuned amplifier provides the same 4-volt output, despite the lower input level available in the supergroup control loop. A gain control potentiometer is provided in the pilot amplifier feedback loop for manually adjusting the output level of the regulator.

The dc amplifier provides the necessary thermistor current by expanding departures in the input signal from the reference voltage. As indicated in Fig. 11, the amplifier consists of two silicon transistors sharing a common emitter resistor, in differential configuration. Satis-

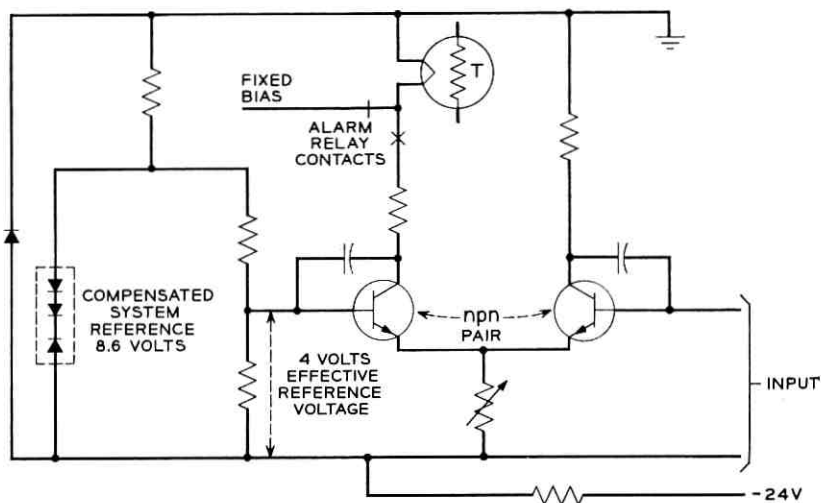


Fig. 11 — Schematic of dc differential amplifier.

factory amplifier stability is provided by paired transistor units in the differential circuit. Since the transistors share a common emitter resistor, collector current is a function of the difference between the input and the reference voltages.

The reference diode (about 8.5 volts) consists of three separate *p-n* silicon junctions, one reverse-biased. Forward-biased junctions are selected to compensate for the thermal characteristics of the reverse-biased junction, providing stability of 0.005 per cent per degree F. The complete dc amplifier contributes about 0.1 db change in regulator level over the range 40°F to 140°F.

The envelope expansion provided by the differential amplifier is a function of the effective reference voltage as well as transistor gain. A reference voltage of about 4 volts provided more than the required 20 to 1 expansion for the gain range of the transistor. The reference voltage was effectively reduced to 4 volts by shunting the reference diode with precision resistors. This reduced the pilot amplifier gain requirements.

Typical operation of the regulator in response to a change in the level of input signal may be summarized as follows. Assume a small step-type decrease occurs in the input signal. This step is transmitted without delay through the transmission amplifier and through the control circuit to the input of the direct current amplifier. The direct-current amplifier immediately responds to the step-type decrease in its input by increasing the current through the thermistor heater. Because of the direct-current amplifier expansion characteristic, the change in thermistor heater current will represent a much greater potential change in the gain of the transmission amplifier. The thermistor begins to change its bead resistance but, because of its thermal inertia, the change cannot be made immediately. As the gain of the transmission amplifier changes, the input to the direct-current amplifier is increased toward its former value. Thus the thermistor current is reduced to a value which represents the change in gain of the transmission amplifier. When equilibrium is reached, a small servo-type error signal remains in the regulator output. This same error signal at the input of the direct-current amplifier causes the required change in amplifier gain that compensates for the original decrease in signal level.

The thermistor provides sufficient delay for a smooth, exponential-type response free from excessive overshoot or hunting. Thermistor time response is, of course, speeded up by the closed-loop expansion. The transient response for small changes is completed in two to three seconds; for large non-linear changes, ten to fifteen seconds. This contrasts with the 60-second open-loop time response of the thermistor.

The intended operating range of the supergroup regulator is ± 6 db from normal level; that of the group ± 4 db from normal. A typical regulation characteristic is shown in Fig. 12. The lower limit is readily adjustable by the maximum thermistor current delivered by the direct current amplifier. The upper limit reflects zero thermistor heater current or minimum transmission amplifier gain.

3.3 Tandem Operation

Tests were conducted on 9 supergroup regulated amplifiers of the final design in tandem. The frequency response was obtained by modulating the pilot at the input to the first regulator and recording the envelope response at the output of the last (ninth) unit. The resulting data is plotted in Fig. 13. It will be noted that a peak gain enhancement, limited to about 3 db, occurs at about two cycles per second. The comparable time response for 9 in tandem to a small step disturbance is shown in Fig. 14. This response is compared, for reference, with that of a completely damped system. Overshoot is about 50 per cent, which is approximately that predicted by the model (Fig. 10).

Five group regulators in tandem measured about 3.5 db of gain enhancement, with time response producing overshoot 50 per cent of the original step. The additional gain enhancement of the group circuit is due to the sharper cutoff characteristics of the group pilot bandpass filter.

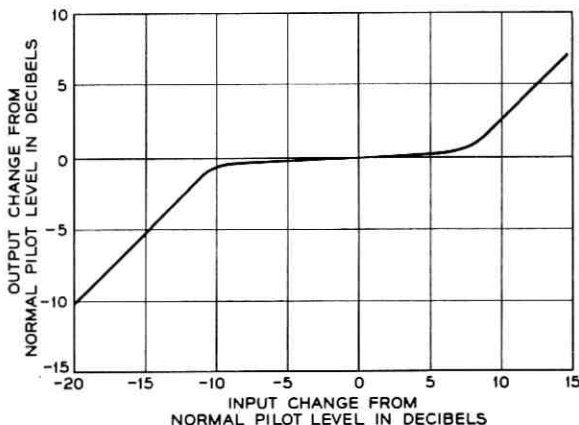


Fig. 12 — Supergroup regulator — measured regulation characteristic.

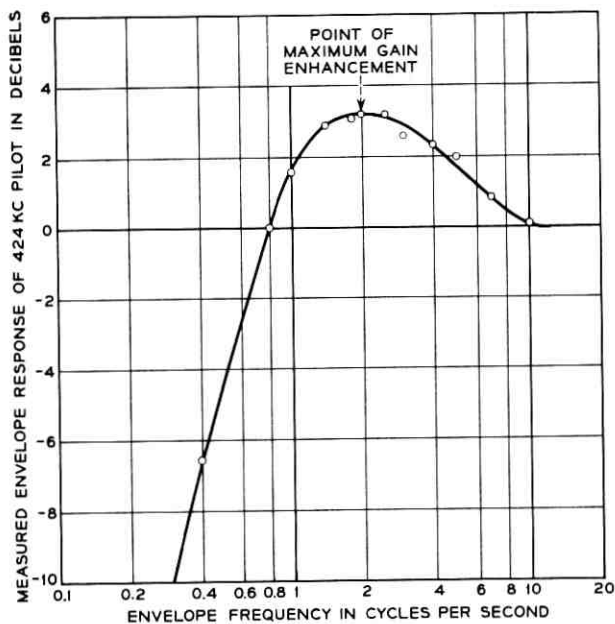


Fig. 13 — Measured frequency response of nine supergroup regulated amplifiers in tandem.

3.4 Trial Results

Several experimental regulator circuits were installed in L1 receiving terminals to gain field experience. Later, a more extensive trial was undertaken in which 50 prototype regulators and amplifiers, 20 supergroup and 30 group, were installed in terminals covering a wide geographical area from Chicago to California.

Performance data was obtained over a one-year period using automatic trunk testing equipment which sets up connections and measures voice-channel net loss on an automated basis. The data indicated substantially improved net loss stability, and led to the decision to incorporate regulation in the L600 and L1860 terminals.

IV. OTHER TRANSMISSION CIRCUITS

4.1 Amplifiers

Several transistor feedback amplifiers have been developed for use in the new terminals. These fall in two main categories: transmission

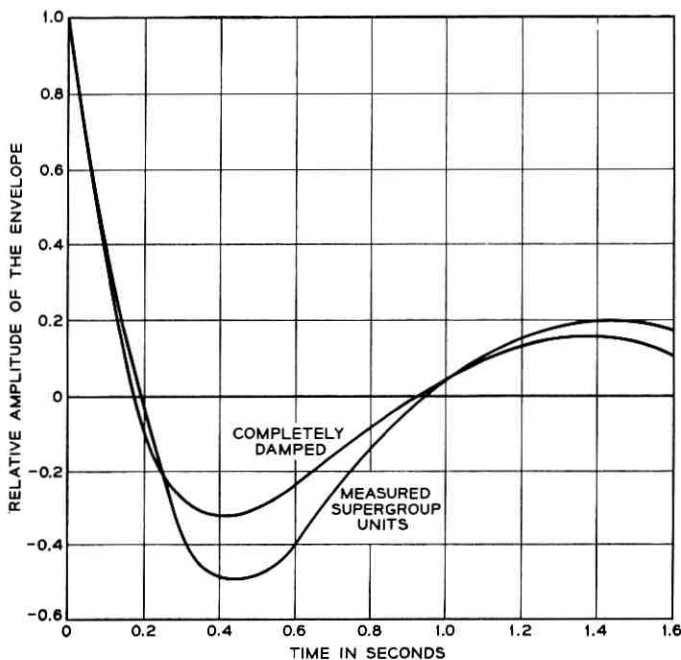


Fig. 14 — Measured time response of nine supergroup regulated amplifiers in tandem vs that of a completely damped system.

amplifiers and pilot amplifiers. The transmission amplifiers are designed for constant gain in the band of interest, resistive input and output impedances to match the external circuit, and power capabilities to handle the expected loading with low nonlinear distortion. The pilot amplifiers are designed for large stable gains at one frequency. Since there is much similarity among the various types in each category, only one of each will be described.

Diffused silicon npn transistors are used in all amplifiers. This type was chosen because of its adequate high frequency performance for this application, the permissible junction temperature which is higher than germanium transistors, and its anticipated low cost in quantity production.

Printed wiring boards are used for all amplifiers to reduce variations in wiring and assembly, insuring good reproducibility. Input and output transformers have manganese-zinc ferrite cores which have been developed to yield the same performance as obtained with earlier designs requiring up to 100 times as much volume. The transformers were designed for mounting directly on printed wiring boards.

4.1.1 *Receiving Group Amplifier*

Each receiving group requires an amplifier with a nominal gain of 38 db. The gain is adjusted by the pilot-controlled regulator over a range of ± 4 db from nominal. The transmission level at the output is -5 db, referred to zero transmission level. Good linearity is required at levels up to the peak loading of 12 voice channels.

The schematic is shown on Fig. 15. Hybrid bridge feedback is used at both the input and output to control the impedances. Local emitter feedback is provided on the second and third stage to improve the linearity at high output power, and to control the maximum value of loop gain. A silicon junction voltage-limiter diode is used to stabilize the bias circuit. The supply voltage is nominally 24 volts dc.

Two equal outputs are obtained from the output hybrid transformer. One is the regular output, and the second output is divided between the pilot-controlled regulator, and the group test jack used for monitoring.

As described earlier, the regulator controls the gain of the amplifier by controlling the direct current supplied to the heater of an indirectly heated thermistor. The thermistor element resistance forms the shunt portion of the feedback network.

The output transistor has a collector dissipation of about 0.35 watts. To obtain a junction temperature which would not jeopardize long life, a 23A transistor is used which has a fluted heat radiator attached to the case. A production model of the amplifier is shown in Fig. 16, which also shows the pilot amplifier described in the next section.

The gain characteristics are shown on Fig. 17. The forward gain in the 60–108-ke band is over 75 db, making the minimum feedback about 35 db. Measurement of the loop gain and phase shift characteristics indicate that a phase margin of at least 30° and a gain margin of 10 db was realized.

4.1.2 *Pilot Amplifier*

The schematic of the pilot amplifier for the 92-ke group pilot is shown in Fig. 18. This amplifier provides about 47 db of power gain to raise the pilot signal to the level needed, after rectification, to drive the dc amplifier in the group regulator. A second output is used to supply the pilot alarm circuit. The antiresonant circuit in the feedback path reduces the gain more than 25 db at frequencies more than 30 ke from the pilot.

The gain control has a range of 6 db and is used to adjust the output level of the regulated group amplifier. Since the pilot amplifier is in the control loop of the regulator, gain variations in it would directly affect

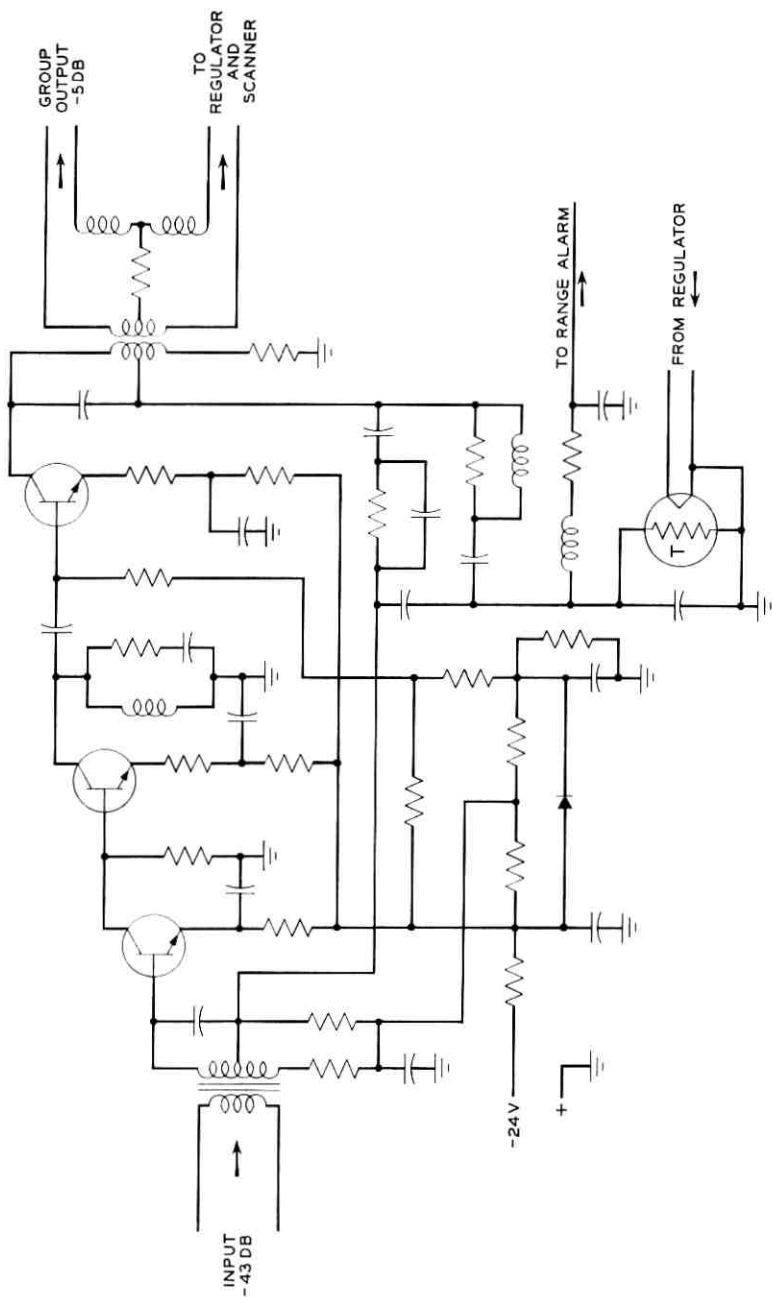


Fig. 15 — Schematic of receiving group amplifier.

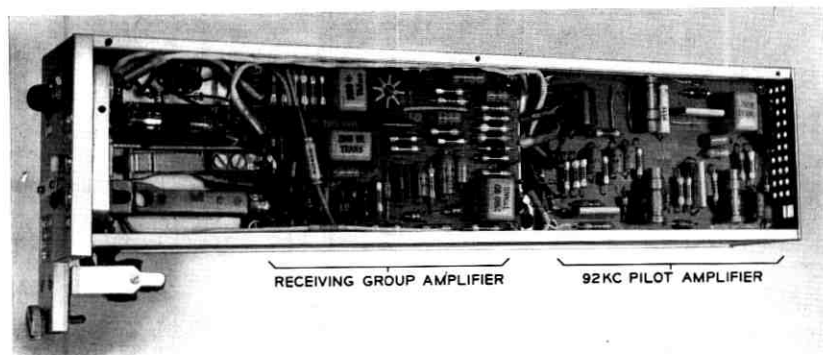


Fig. 16 — Receiving group module with side cover removed: monitoring jacks and thermistor on left, group amplifier in center, and 92-ke pilot amplifier on right.

the output level. Variations are reduced by over 35 db of feedback at the pilot frequency, and by local feedback on the first two transistors. The loop gain and phase shift characteristics are shown on Fig. 19. The feedback circuit from collector to base of the third stage is antiresonant at the pilot frequency, increasing the forward gain at 92 ke, but reducing it at other frequencies at the same rate that the feedback circuit reduces the amplifier gain.

4.1.3 Computer Program

A special computer program¹⁰ has been developed for computing the performance of these amplifiers. It serves as an aid to the design and to indicate the effect of variations in parameters on the performance. The program solves up to 23 nodal equations representing the amplifier parameters and schematic. Subsidiary programs are used to convert transistor and transformer measurements to the hybrid parameters used in the program.

4.2 Modulators

4.2.1 Choice of Element

The modulators and demodulators in the earlier group banks and in the lowest 8 supergroups employed copper-oxide varistors for the modulating devices.¹¹ These have some obvious disadvantages such as high shunt capacitance, requiring a low-impedance circuit for their use in group and supergroup modulators. This in turn leads to instability with temperature and aging, causing small deviations in loss. However, the

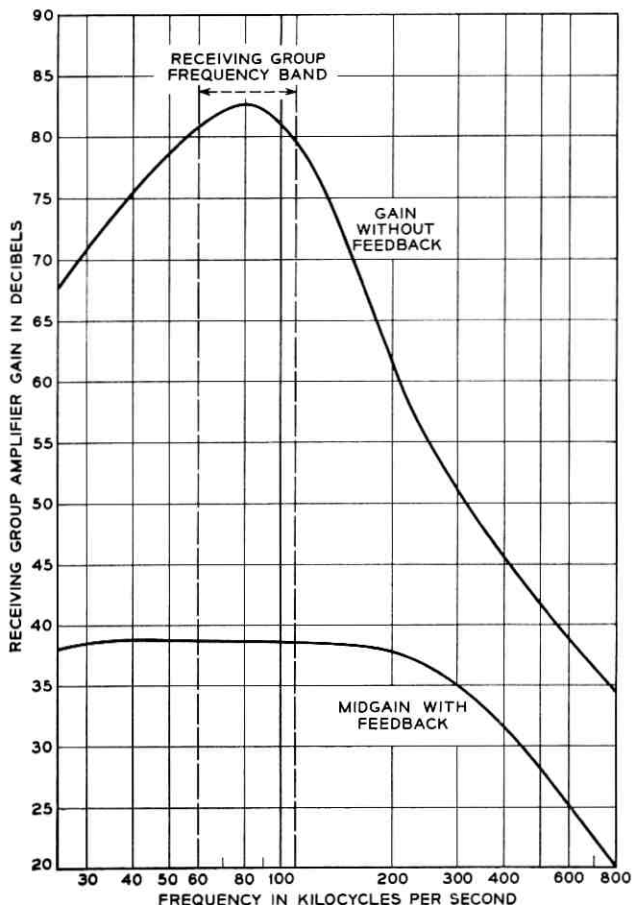


Fig. 17 — Receiving group amplifier gain characteristics.

device has low noise, produces low harmonic distortion, and has performed satisfactorily in general. With the increased use of silicon devices in applications formerly filled by copper oxide, the demand for copper-oxide devices is decreasing and it may not be long before it is uneconomical to provide them for modulators.

With these considerations in mind and after preliminary tests of silicon diodes, it was decided to use silicon diodes in the modulators for the new terminals.

An examination of available types of diffused junction silicon diodes led to the selection of the 432A diode. This diode has less than 4 pf

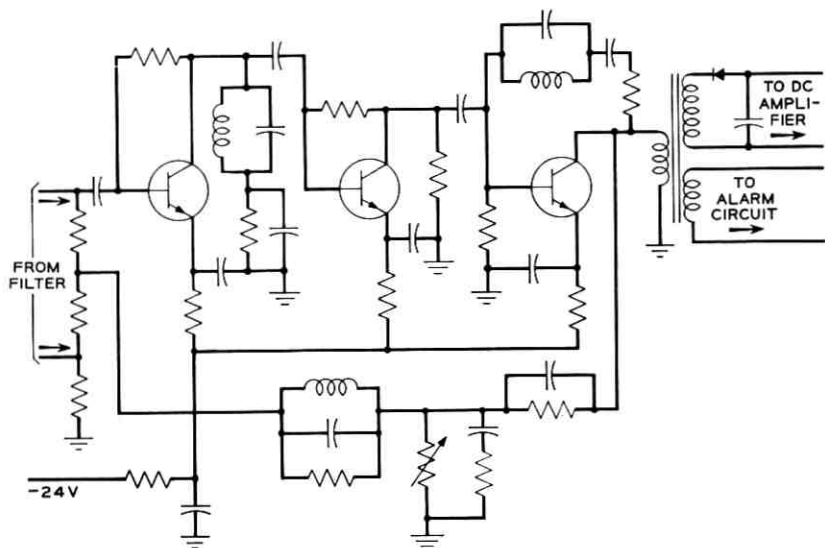


Fig. 18 — Schematic of 92-kc pilot amplifier.

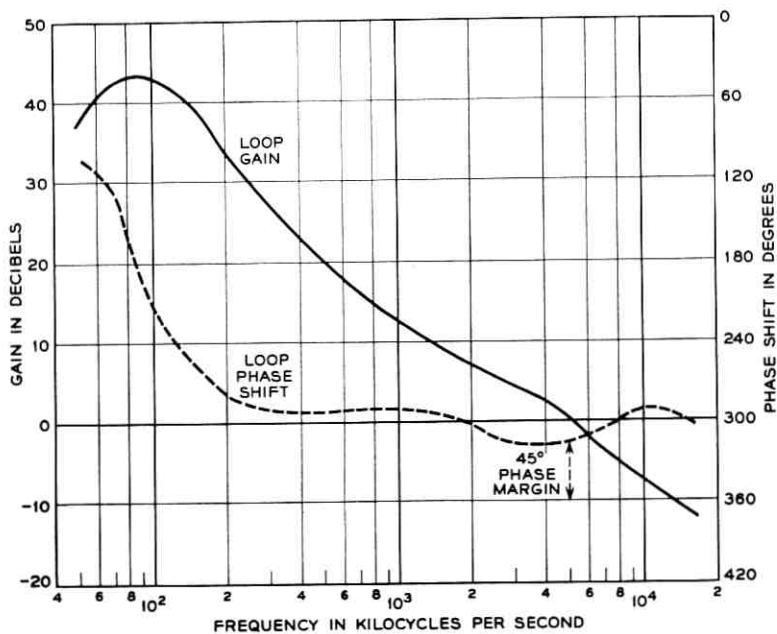


Fig. 19 — Pilot amplifier loop characteristics.

reverse capacitance and a recovery time of less than 4 nanoseconds. With a forward current of 10 milliamperes, the ac resistance is less than 20 ohms. The characteristics are sufficiently uniform so that it is economical to select quads having carrier balances of greater than 35 db.

4.2.2 Modulator Performance

When used in the conventional ring modulator circuit shown on Fig. 20, the diode has yielded satisfactory performance for both group and supergroup modulation and demodulation stages. An impedance level of 1,000 ohms is used to terminate the ring section. This makes the insertion loss of the diodes almost negligible when the carrier level is sufficient to produce a forward voltage of 0.8 volt. The carrier is supplied at an impedance level of 40 ohms.

A group carrier level of +15 dbm is standard on existing terminals and this has been continued on the new ones. With the silicon diodes, the level can be decreased to 10 dbm or increased to 22 dbm before the modulator loss to the wanted sideband increases 0.2 db. Reducing the level below +14 dbm increases the harmonic distortion to an undesirable degree. Noise tests have shown the silicon modulators to have a noise figure about 2 db larger than the copper-oxide modulators, but the conversion loss is substantially less.

When the generator and load impedances terminating the ring and its transformers are resistive, the modulator loss is constant with frequency over the group and supergroup bands. When one of these terminations is a bandpass filter, the impedance departs from an ideal resistance, especially due to the effect of the filter impedance at the frequency of

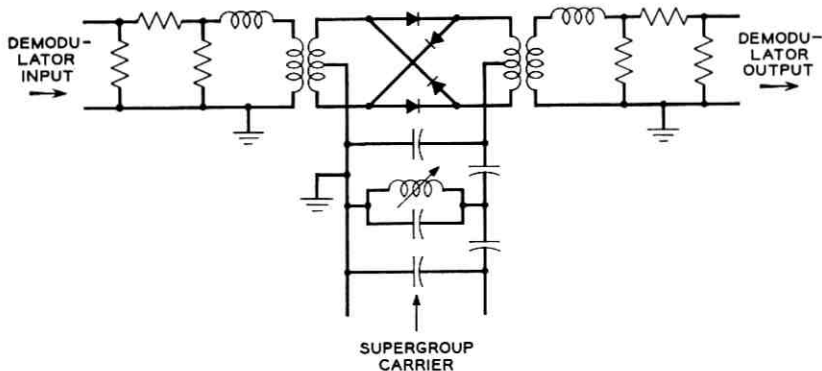


Fig. 20 — Schematic of supergroup modulator or demodulator.

the unwanted sideband. Without impedance correction this introduces loss distortion in the passband of the wanted sideband. The inductors and resistors shown on Fig. 20 perform the necessary impedance matching for the supergroup modulators and demodulators. A similar circuit is used in the group modulators and demodulators.

The final design of modulator circuits and bandpass filters together have passband loss distortion substantially equal to that of the filters alone with resistive terminations.

4.3 *Filters*

A study of the original L-1 multiplex terminal³ showed that over 50 per cent of the space occupied by the terminal was required for filters to separate the frequency spectrum, select carriers and pilot frequencies. Therefore, to achieve a new equipment arrangement it was essential that development work on the filters be undertaken to substantially reduce the size, and if possible, the cost. However, a reduction in the quality of the system to achieve a size and cost reduction could not be tolerated; that is, the same discrimination and distortion requirements were imposed on the new filters.

In the original filters, for example, slug-tuned air core inductors with large molded silver mica capacitors were used. Also, a method of double shielding in the series arm of the filters was incorporated which placed the parasitic capacities to ground across specific shunt branches of the structure. These large air core inductors with sufficient space about them to give a good stable low value of capacitance to ground made the filters quite large. These filters proved to be very stable with temperature and aging and produced practically no modulation in the system.

To obtain a substantial size reduction it became evident at the start that all component apparatus had to be miniaturized, particularly the large air core inductors. Also, the use of double shielding should be eliminated, if possible.

Of the many types of filters required in the terminals only a few typical ones are described in the following sections to illustrate the trend in components, method of assembly and shielding.

4.3.1 *Components*

An investigation of available component apparatus with satisfactory tolerance and size revealed that only capacitors and resistors were available in production for miniaturization of filters. A development

program had to be undertaken on crystal units, inductors and transformers to realize size reductions that would be comparable with capacitors and resistors.

Several types of small ferrite core adjustable inductors have been developed which have good temperature coefficient characteristics, low dissipation factors and have an adjustment range of several per cent. The ferrite core inductors do introduce measurable modulation in the system but it has a negligible effect on the over-all performance. Several small-size ferrite transformers have been developed for use in amplifiers, modulators and filters in the system.

Several types of small size shielded crystal units were developed for use in filters. The units were required to have a low dissipation (high-Q) and good temperature coefficients. A method was developed for encapsulating the quartz crystal plate in an evacuated cold welded container, eliminating the use of glass containers and additional shielding.

A comparison of the old and new style components used in the fabrication of typical filters is shown in Fig. 21. One can readily see that with such a size reduction in component apparatus with substantially the same electrical characteristics, the filter designer had an excellent chance to make sizable reductions in volume.

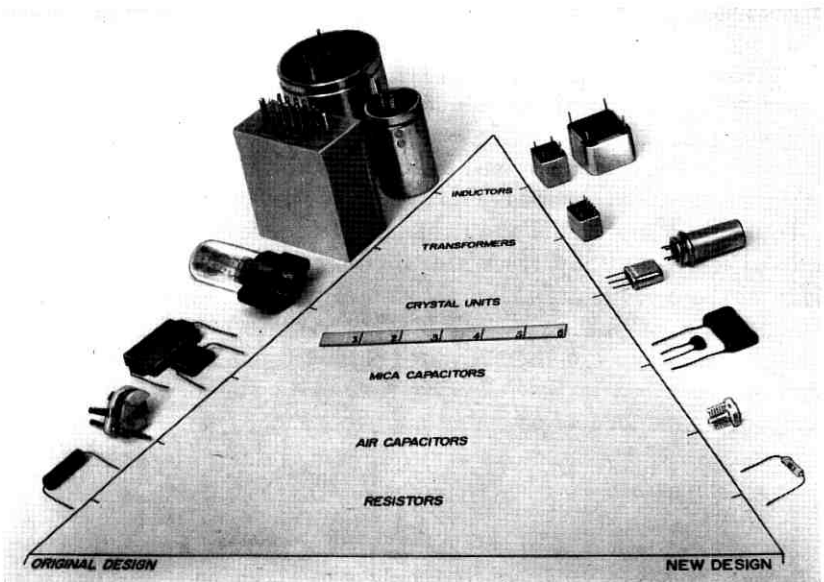


Fig. 21 — Comparison of old and new style components.

4.3.2 Group Bandpass Filters

With component apparatus available that can be mounted by its leads, it was decided after investigating various methods of assembly and wiring to use printed wiring techniques. Printed wiring provides good uniformity and consistent capacitance between leads and to ground, from components in production. Also, with printed wiring it was found possible to obtain low ground resistance paths and satisfactory shielding between various portions of a filter using shielded inductors. To provide adequate shielding from the filters to various other parts of the terminal the filters were inserted in drawn metal containers with soldered covers.

In packaging filters for the system the number of different sizes was kept at a minimum. Also, an arrangement was made to provide some of the filters with a moisture-resistant seal where necessary.

The five group bandpass filters were designed to operate in a 75-ohm unbalanced circuit with alternate filters paralleled, i.e., numbers 1, 3 and 5 in one group and numbers 2 and 4 in the other. Each group at its paralleled side was connected to the opposite port of a hybrid coil with an impedance compensating network shunted across each group to improve the impedance at the line and modem side of the filter. Each filter passes a 48-kc band, the five bands extending from 312 kc to 552 kc.

The schematic of the filter, together with the insertion loss characteristic of a typical group filter, is shown in Fig. 22. This characteristic is adjusted in a visual test set by tuning the slugs in the inductors of the filter under test until it is identical to a reference filter. The volume reduction of this filter over the old design was 30 to 1 and the cost reduction about 3 to 1. Each of the filters and the compensating networks plus hybrid coils were assembled in a metal container $1\frac{1}{2} \times 2\frac{11}{16} \times 4\frac{11}{16}$, exclusive of terminals and studs. These six units were assembled on the rear of the group bank equipment shelf and wired together with small coaxial cables. Fig. 23 shows the internal arrangement and external appearance of the filter.

The performance of this new miniature filter was similar to the original larger filter except for modulation, which was poorer but entirely satisfactory for system performance.

4.3.3 Supergroup Bandpass Filters

The new supergroup bandpass filters with the exception of Supergroups 1 and 3 are identical in size and schematic to the group band filters. All supergroup band filters have a band width of 240 kc. The

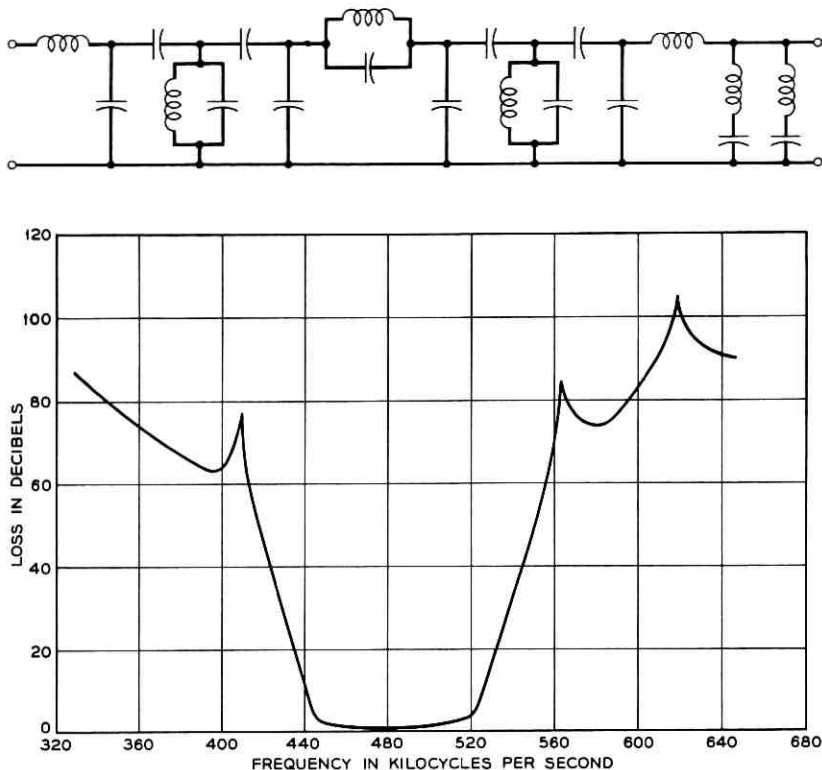


Fig. 22 — Schematic and insertion loss of a typical group band filter.

ten supergroup bandpass filters were designed to operate between 75-ohm impedances with alternate filters paralleled, numbers 1, 3, 5, 7 and 9 in one group and 2, 4, 6, 8 and 10 in the other. Each group at their paralleled side is connected to opposite ports of a hybrid coil with a compensating network shunted across each group for impedance and distortion improvement. In order to meet the discrimination requirements for the high-frequency filters it was necessary to use a special grounding arrangement. Also, to meet the crosstalk requirement between the upper supergroup filters it was necessary to use small coaxial jacks instead of terminals on the filters with all connections made with coaxial cabling. Fig. 24 shows a typical schematic and insertion loss characteristic of a supergroup filter except 1 and 3, connected in parallel with other supergroup filters.

In the frequency range above 2 mc it has not been possible to obtain as good a distortion characteristic across the useful band of 240 kc with

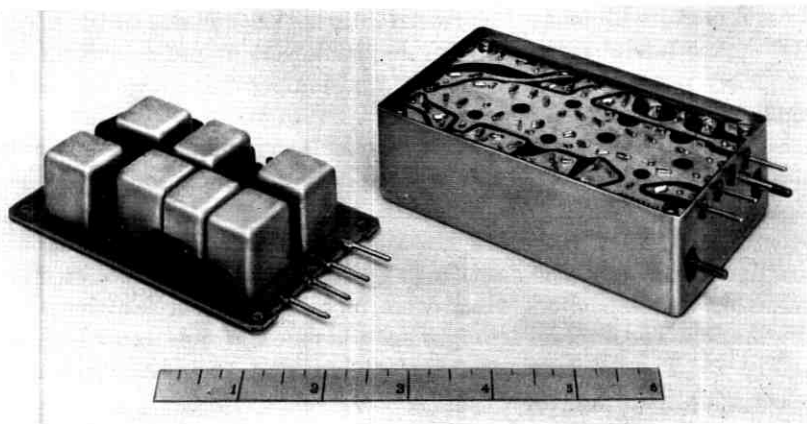


Fig. 23 — Mechanical construction of group band filter, right; component side of board, left.

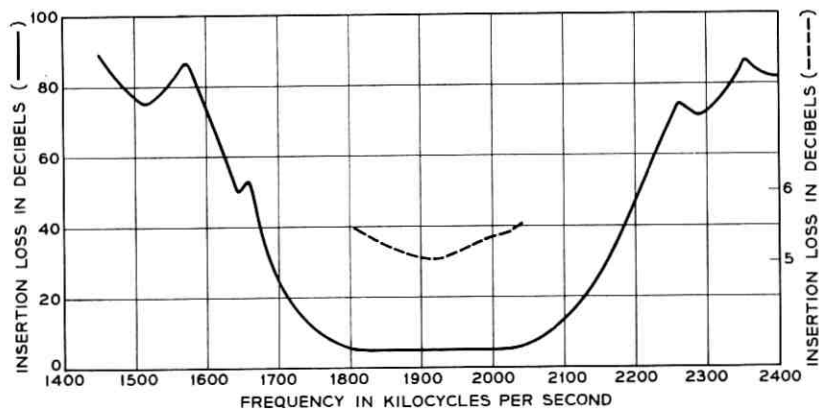
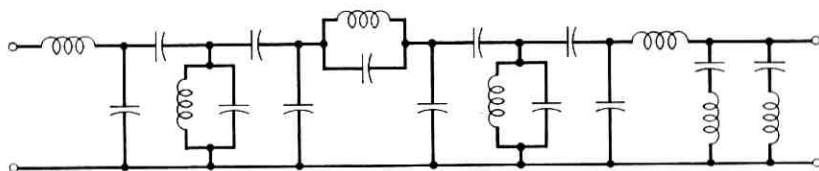


Fig. 24 — Schematic and insertion loss of typical supergroup band filter.

these miniature filters as the original filters. This was due to the poorer "Q" in the ferrite inductors. Development work is progressing in this frequency range to improve the "Q" and temperature coefficient of the small ferrite inductors.

Supergroup filters number 1 and number 3 were designed as minimum inductor asymmetrical LC filters with sharp discrimination on one side of each of the filters. This was necessary to attain a high discrimination over the frequency range 312 kc to 552 kc. These filters are larger in size to accommodate the larger high-Q ferrite inductors and a few more inductors and capacitors which were required to meet the sharp discrimination. The internal arrangement and external appearance is shown in Fig. 25. The schematic and typical insertion loss characteristic of Supergroup No. 3 are shown in Fig. 26.

To equalize the distortion across supergroup bands 1 and 3 to less than 0.5 db, equalizers are provided which insert the correction at corresponding frequencies in the basic supergroup band, 312-552 kc.

4.3.4 *Crystal Filters*

In practically every communication system, extremely narrow band elimination and bandpass filters are needed to remove extraneous noise or frequencies from a spectrum so that pilot frequencies may be inserted and picked-off further along in the system to control gain or shaping. In this system crystal filters were used for this purpose because of their extremely high Q, good temperature characteristics and stability.

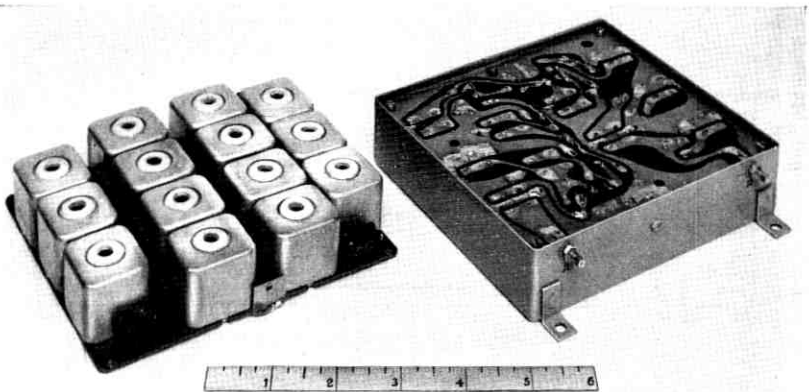


Fig. 25 — Mechanical construction of supergroup filter no. 3: right, printed wiring side; left, component side.

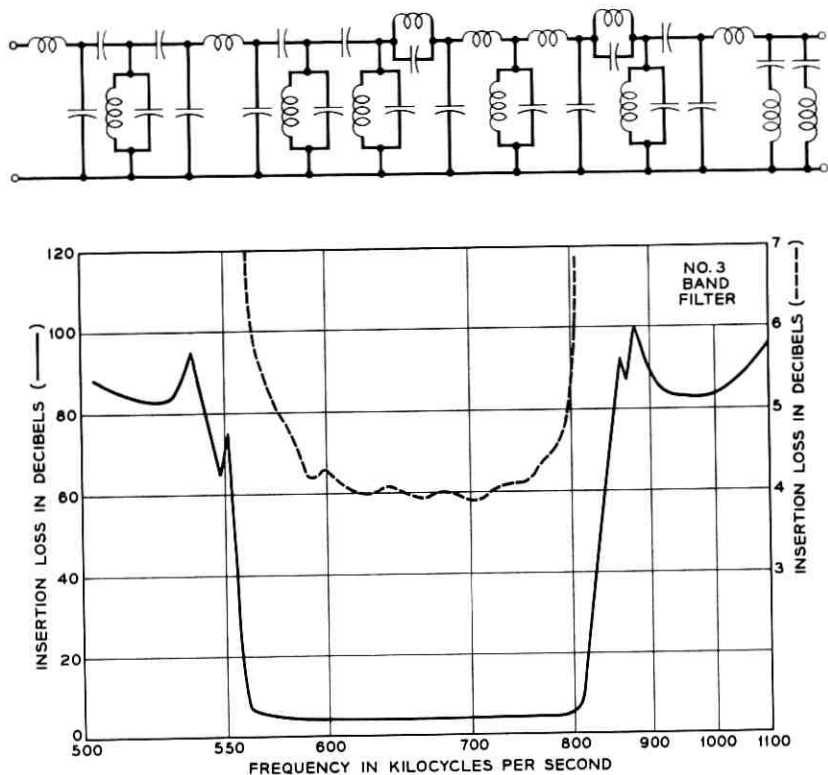


Fig. 26 — Schematic and insertion loss characteristic of supergroup filter no. 3.

Two such typical crystal filters are the 92-kc band elimination and bandpass filter. The 92-kc band elimination had to suppress a narrow band about 92 kc, pass the voice-frequency channels either side of 92 kc with practically no distortion, have a good return loss over the 60 to 108-kc band and operate over the central office temperature range. The schematic of filter which was found to be most economical for this job together with its insertion loss characteristic is shown in Fig. 27.

The schematic of the crystal bandpass filter which selects this 92-kc pilot frequency is shown in Fig. 28 with its typical insertion loss. It was designed as a lattice section to operate from a 135-ohm balanced impedance to a 1,000-ohm unbalanced impedance.

The internal arrangement and external appearance of these filters are shown in Fig. 29.

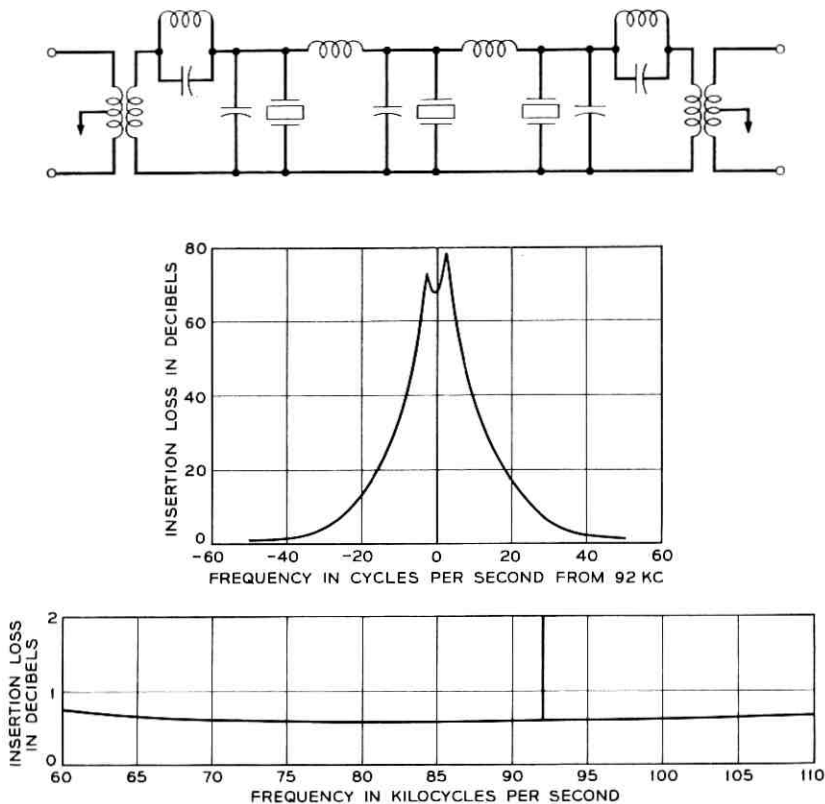


Fig. 27 — Schematic and insertion loss of the 92-kc band elimination filter.

4.3.5 Carrier Supply Filters

In the carrier supply portion of the terminals only "LC" filters are used because of the load carrying requirements and the availability of high-Q ferrite inductors for those moderately narrow band filters. Also, from cost consideration it proved to be more economical to use LC rather than crystal filters.

The channel carrier supply filter has to operate in parallel at the output of the harmonic generator and provide 45-db and 75-db discrimination 4 kc and 8 kc respectively, either side of the selective harmonic. The schematic of the filter used to meet this requirement with a typical loss characteristic is shown in Fig. 30.

The group carrier supply filters connected to the output of a 12-kc harmonic generator have to meet requirements similar to the channel

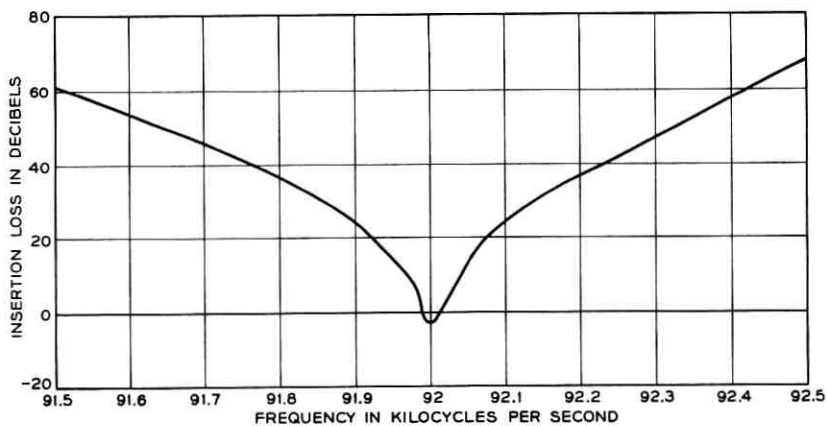
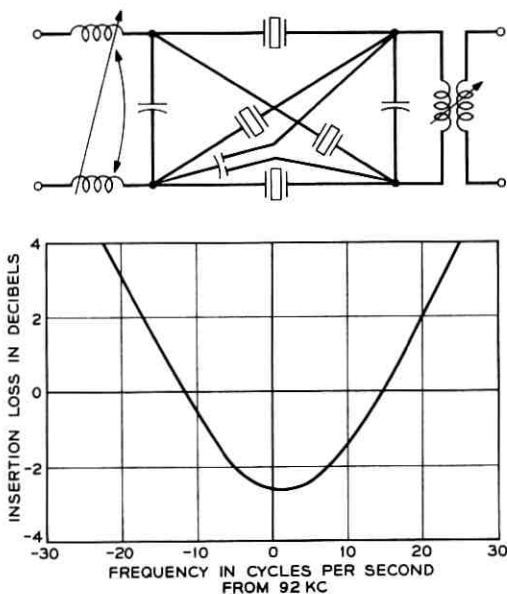


Fig. 28 — Schematic and insertion loss characteristic of the 92-kc crystal bandpass filter.

carrier supply filters in a higher frequency range. The schematic of these filters together with a typical loss characteristic is shown in Fig. 31.

The filters for the supergroup carrier circuit were constructed as two identical individual filters with a transistor amplifier between them.

All of these filters were constructed with the same type of components in the same manner as the other filters of this system.

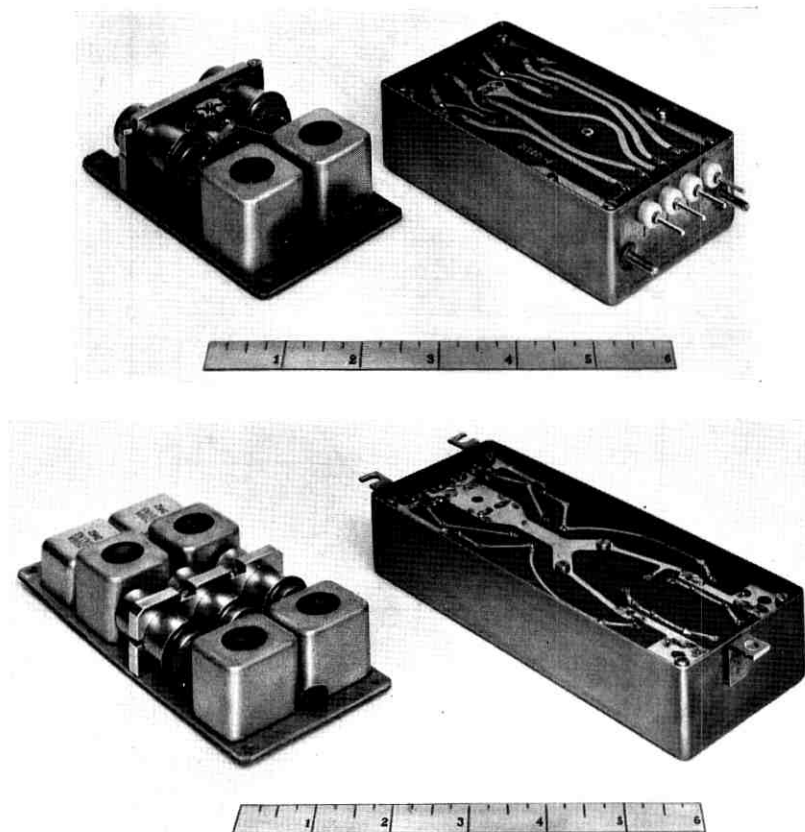


Fig. 29 — Mechanical construction of the 92-ke bandpass filter (upper photo) and elimination filter (lower photo). Component side of board at left, printed wiring side at right.

4.3.6 Group Connectors

At branching points it is frequently required to transfer a group of 12 channels from a receiving terminal to another transmitting terminal. This can be done at the basic group frequency, 60–108 ke, by the use of a bandpass filter. This filter must furnish the discrimination needed to suppress portions of adjacent groups which are not suppressed by the group bandpass filters.

A new group connector, the B-2, has been developed with less attenuation distortion than that of the earlier design, now known as the B-1. Discrimination of more than 70 db is provided at frequencies below 59.7 ke and above 108.6 ke. A typical passband and discrimination frequency

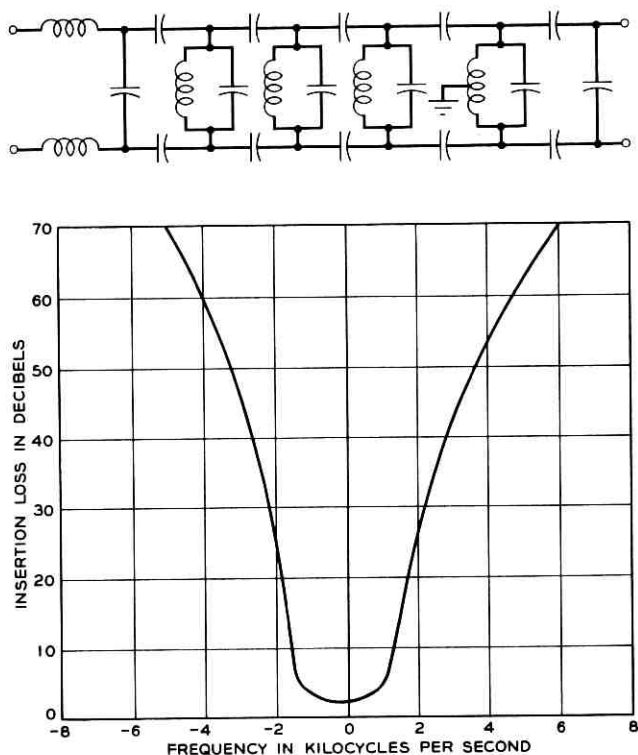


Fig. 30 — Schematic and insertion loss of channel carrier supply filter.

characteristic is shown on Fig. 32. The envelope delay distortion in the band, 64–104 kc, is about 100 microseconds. For applications such as data transmission requiring constant delay, a delay equalizer can be provided.

The filter used in the connector employs special wide crystal filter sections which give a high degree of selectivity at the band edge. The schematic of this filter is shown in Fig. 33.

A model of the new connector is shown on Fig. 34. In addition to the filter, an attenuator is supplied to adjust the loss of the connecting circuit, and space is available for adding a delay equalizer.

4.4 *Supergroup Connector*

When it is desired to transfer an entire supergroup from a receiving terminal to another transmitting terminal, this can be done at the basic

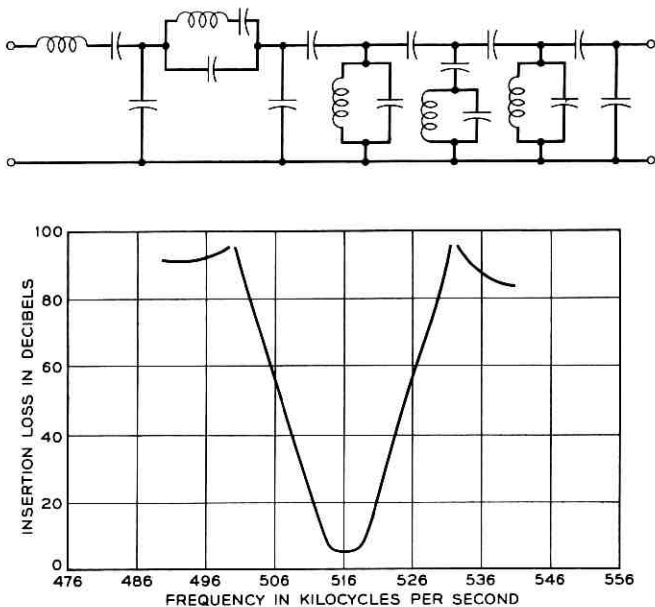


Fig. 31 — Schematic and insertion loss of group carrier supply filter.

supergroup band, 312–552 kc, with the use of filters and an amplifier. This equipment is known as the C-2 supergroup connector, replacing the earlier C-1 version. A block schematic is shown on Fig. 35.

The bandpass filter has been designed as a minimum inductor structure using high-Q ferrite inductors and has over 80-db discrimination at frequencies below 304 kc and above 560 kc. The passband distortion is less than ± 0.2 db. A delay equalizer is being developed to equalize the delay distortion of this filter for data transmission. When this is not required it can be replaced by a pad.

Two pilots at 308 and 556 kc, respectively, occur in certain supergroups on certain systems. To prevent interference, these are suppressed by a crystal band elimination filter. Crystal units are used because of the high degree of selectivity required and are mounted in a temperature controlled oven.

A transistor amplifier provides the gain needed to compensate for the loss of the filters and equalizer. The transmitting group intermediate amplifier and associated low-pass filter are used for this purpose.

The filters and equalizer preceding the amplifier have been designed to occupy the space in the receiving bay which is provided for the five re-

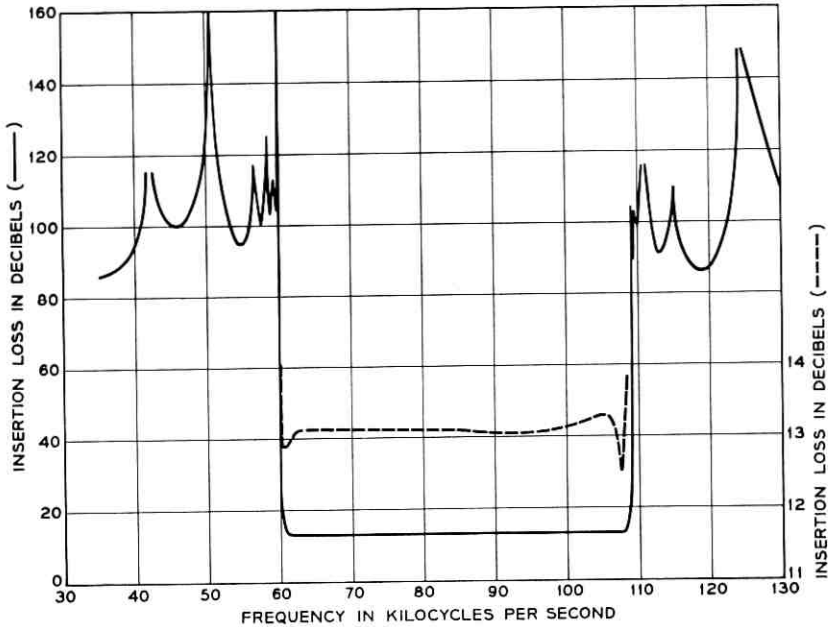


Fig. 32 — Insertion loss of the B-2 group connector filter.

ceiving group amplifiers. When a supergroup connector is used, these are not required, and the supergroup connector filters may be installed as shown on Fig. 36, without modifying the bay wiring.

The amplifier and low-pass filter are located in the transmitting terminal in place of the corresponding equipment for the transmitting group bank.

V. AUXILIARY CIRCUITS

5.1 Alarm Circuits

Several types of alarm circuits are provided to alert office personnel to conditions requiring their attention. Three of these indicate a service interruption: fuse failure, carrier supply, and loss-of-pilot. A fourth indicates when a regulator nears the end of its effective range. This may indicate an incipient trouble condition.

The fuse alarm is a standard circuit in which a blown fuse closes the alarm circuit through a bus provided for this purpose. The carrier supply alarms are described in a companion paper.⁴

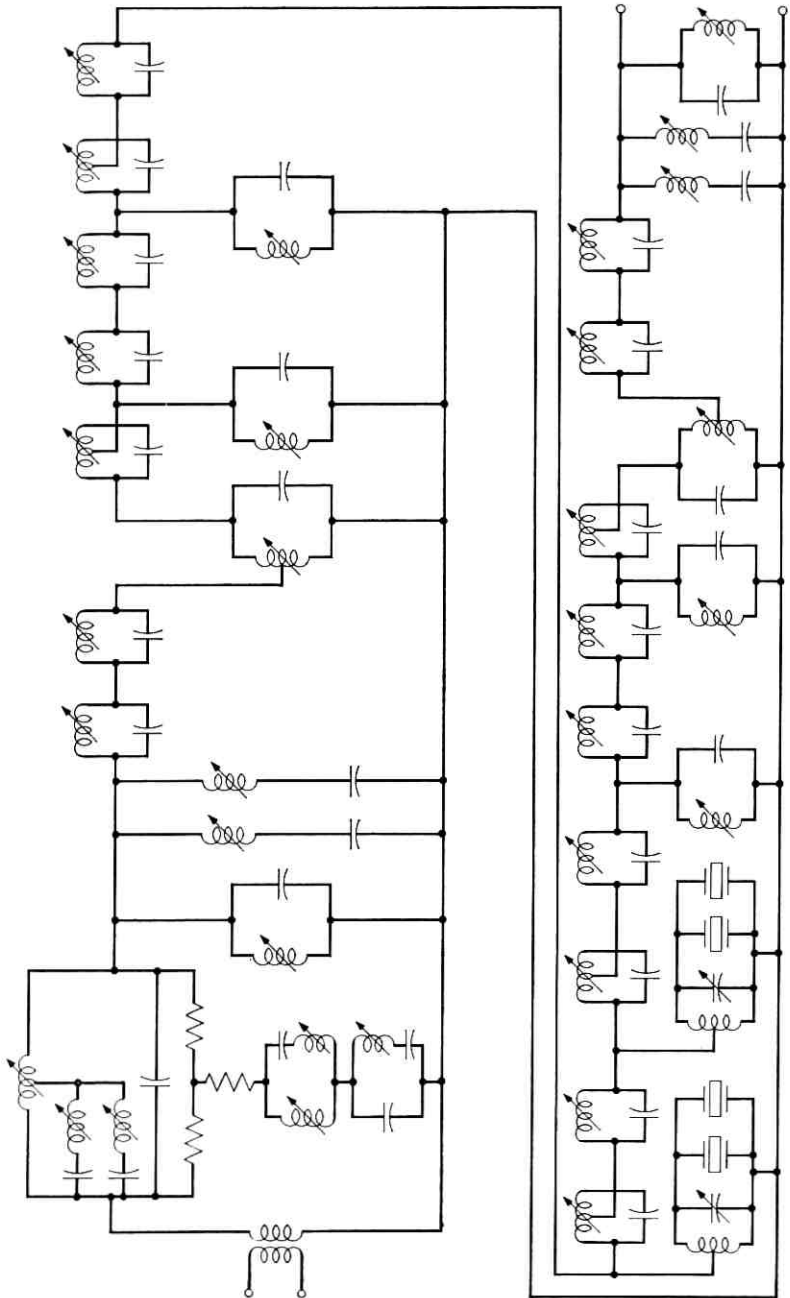


Fig. 33 — Schematic of the group connector filter.

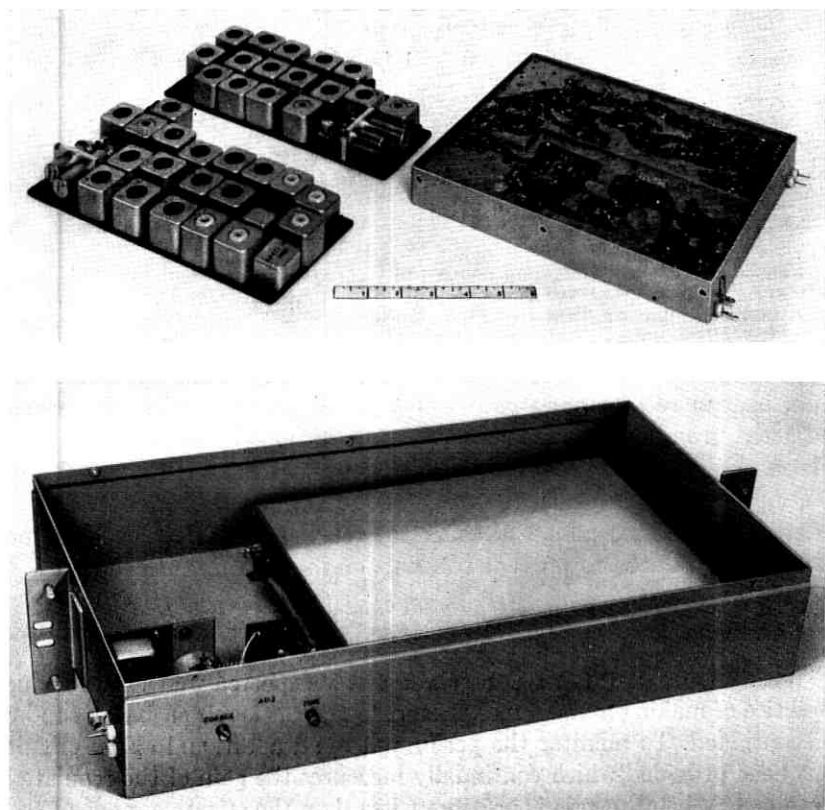


Fig. 34 — B-2 group connector. Upper: internal construction of filter. Lower: assembly with cover removed

5.1.1 Pilot Alarm

The loss-of-pilot alarm indicates that the pilot has disappeared, or become too weak for satisfactory regulation. When this occurs, the operation of a miniature relay in the receiving group or supergroup module transfers the thermistor heater winding to a fixed current source instead

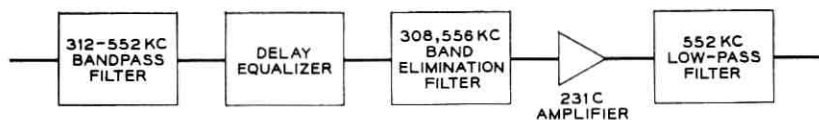


Fig. 35 — Block schematic, C-2 supergroup connector.

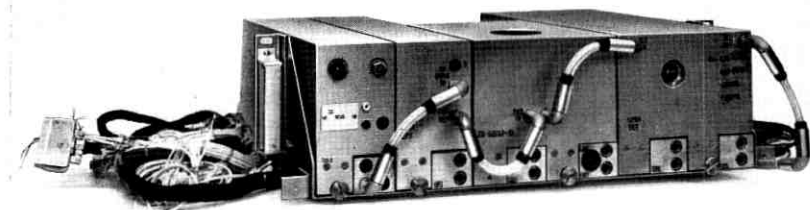


Fig. 36 — Receiving terminal portion of C-2 supergroup connector. Units are, left to right: supergroup receiving amplifier, bandpass filter, pad or delay equalizer, band elimination filter to suppress pilots.

of the regulator. This maintains the amplifier at a mid-range value of gain instead of maximum gain which would interfere with service on adjacent groups or supergroups. The relay also closes a contact to the alarm circuit. A portion of the received pilot current is amplified and rectified and used to hold the relay in the normal, operated condition. A weak pilot or no pilot will cause the relay to transfer. A short time delay is introduced to prevent transfers on momentary interruptions.

5.1.2 *Range Alarm*

With the use of automatic regulation, it is expected that the previous practice of making manual measurements of pilot levels each day can be discontinued. To monitor the performance, an additional alarm circuit has been provided which continually measures the gain of the regulated amplifier. When the gain is greater or less than the effective range of the regulator, an alarm is actuated.

This condition could be caused by malfunction of terminal equipment or by an out-of-limit pilot being received from the transmission system. A sudden loss of pilot will actuate the pilot alarm and prevent the range alarm from occurring. In general, the range alarm indicates an operating condition which is marginal and should be corrected, although service may still be satisfactory.

From its nature, the range alarm lends itself to time-shared operation. It is not essential that the condition be discovered immediately; a delay of several minutes can be tolerated. This makes it feasible to measure the gain of each amplifier in a terminal in turn, repeating the cycle continuously. A scanning circuit, as described in a later section, was developed to facilitate measurements of pilot level. By incorporating means for measuring amplifier gain this scanner has been made to serve both purposes.

The shunt resistance of the amplifier feedback circuit is supplied by the thermistor element. There is a direct correlation between this resistance and the gain of the amplifier. The resistance can be measured conveniently by supplying a small direct current and measuring the voltage. The amplifier portion of the circuit is shown on Fig. 15. Blocking capacitors isolate the measuring circuit. Filtering is also required to prevent noise and crosstalk from entering the amplifier through the measuring circuit.

A differential dc amplifier, the 234B amplifier, has been designed to amplify the dc voltage across the thermistor element. One output of this amplifier drives a meter to display the gain reading. Two additional outputs operate an alarm relay when the gain is outside the effective regulating range.

The connection between the amplifiers and the differential amplifier is by means of the scanner circuit described later. To simplify the scanner, the group amplifier measuring current is transmitted by a simplex system on the group test balanced pair and the shield. At the output of the scanner, the dc signal is separated from the group test voltage by a transformer.

In order to have a means for registering an alarm condition without interrupting the scanning, a magnetic latching relay has been provided for each supergroup shelf. Each shelf accommodates one supergroup amplifier and the associated five group amplifiers. The relay is latched by a pulse from the alarm relay in the 234B amplifier when an end-of-range condition is detected. It remains latched, lighting an alarm lamp on the shelf, until it is reset manually.

5.2 *Pilot Measuring Circuit*

The group pilot at 92 kc and the supergroup pilot at 424 kc are used to control the group and supergroup regulators as described above. They are also used for maintenance, in measuring gains or losses of portions of the equipment, and as an indication of the performance of the transmission system.

A special pilot measuring circuit has been developed to measure these two pilots at the receiving terminal. The operation has been reduced to its simplest form: the signal is patched into the measuring circuit and the deviation from normal level is indicated on a specially calibrated meter.

The group pilot requires a three-stage common-emitter transistor amplifier before rectification. The supergroup pilot is at a lower level and requires an additional two-stage pre-amplifier. All three amplifiers have

considerable feedback to stabilize the gain against variations. The pilot bandpass filters are crystal filters designed to have a stable passband loss, as well as a stable frequency characteristic. These precautions are needed to achieved the desired accuracy.

After rectification, the difference between the rectified pilot and a dc reference voltage is used to drive a meter. The meter displays the deviation from normal pilot level and has a range of ± 1 db.

5.3 Scanner Circuit

A scanner circuit has been developed to provide easy access to the receiving group test outputs and to the second outputs of the supergroup amplifiers. The equipment has been designed to accommodate all the regular and spare amplifiers on three receiving bays (three L600 multiplex terminals or one L1860 multiplex terminal).

The circuit contains wire spring relays arranged to make the following connections: the test output from any selected supergroup amplifier to a supergroup test jack, the test output from one of the associated five group amplifiers to the group test jack, and the corresponding thermistor terminals to the range alarm circuit. Relay walking circuits are included so that the scanning can be automatic, measuring each amplifier in turn in the three bays and then repeating the cycle. About 15 minutes is required for the complete cycle.

Digital display lamps are included in the meter assembly shown on Fig. 37, to indicate which bay, supergroup, and group are being measured. The meters show the deviations from normal for group and supergroup pilot and for the gain of the two amplifiers.

The control panel for the scanner is shown on Fig. 38. Any group or supergroup can be selected manually and held connected, or scanning can be initiated by depressing the proper key.

The pilot meters indicate the pilot deviations when the test jacks are patched to the pilot measuring circuit. Signals at the other frequencies in the group and supergroup bands can also be measured by connecting a suitable measuring circuit to the test jacks.

VI. PERFORMANCE

Extensive tests have been made on several L600A Multiplex transmitting and receiving bays and performance has been at least equal to the objectives in all respects affecting voice transmission and voice channel data transmission. In the rapidly developing field of wideband data transmission, firm objectives for group and supergroup bands have not yet been established. The performance of L1860A multiplex will be tested when the first installation is completed in early 1963.

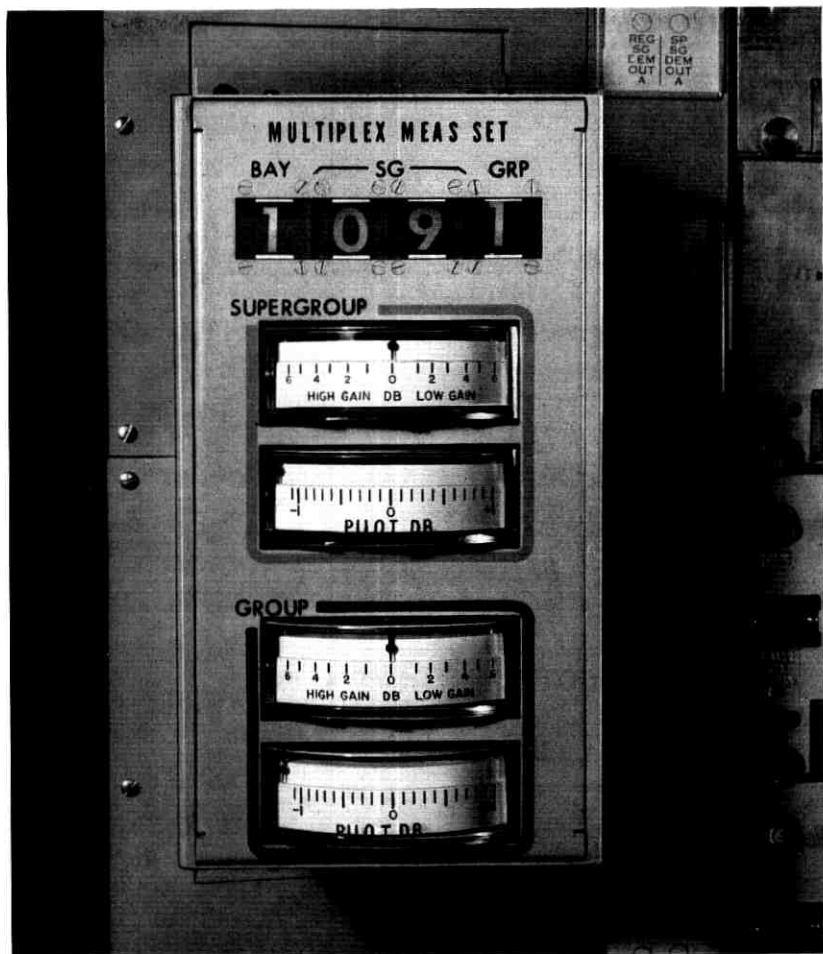


Fig. 37 — Meter display unit for receiving terminal.

The addition of transmitting group amplifiers has resulted in lower noise levels. With a transmitting group bank patched to the receiving bay at the group bank output, and with an A5 channel bank connected to the receiving group output, channel noise, as measured on the 3A Noise Measuring Set,¹² indicated about 15 dbrn at the +7 db level with "C" message weighting (equivalent to 2 dba at zero level). With white noise loading on all channels in the supergroup except three adjoining channels, the noise level in the center quiet channel increased to about 16.5 dbrn when the loading was -15 dbm per channel at zero transmission level.

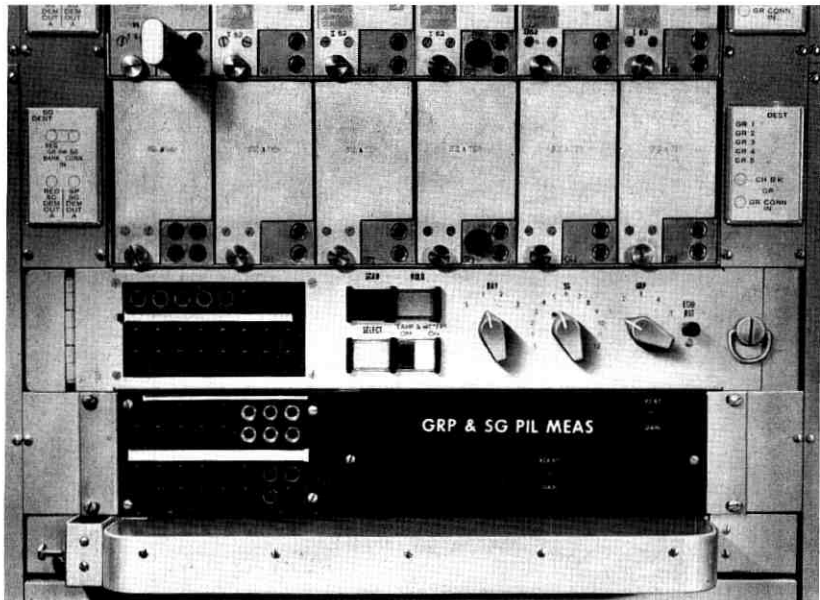


Fig. 38 — Central portion of receiving bay showing unequipped Supergroup No. 1 shelf, scanner control panel, pilot measuring unit, and writing shelf.

Temperature tests on 26 receiving groups equipped with regulators showed that with constant input pilots, the variation in output pilot levels averaged less than ± 0.2 db from 40° to 120° F. Similar tests on 8 supergroup receiving amplifiers showed about twice as much variation. When the input pilot levels were varied ± 4 db for the groups, the outputs remained within ± 0.25 db. For ± 6 db variation in supergroup input levels, the outputs changed less than ± 0.35 db.

All equipment on both the transmitting and receiving bays was designed to operate from 24 volts dc. Tests showed satisfactory operation with supply voltages from 22 to 26 volts, and marginal operation from 20 to 28 volts. At higher voltages, some parts of the circuit overheat and reliability will be decreased. Below 20 volts, some of the relays in the scanner and alarm circuits did not operate.

VII. EQUIPMENT FEATURES

One of the major problems that has been created by the large use of the L1 and L3 carrier telephone terminals is that of interbay cabling. In many offices, the available overhead cable racks have become overloaded, forcing the addition of more racks where possible. Some of the

congestion has been caused by changes in circuit assignment of the equipment. Generally this reassignment required new cabling and when the new cables were run, the old cables were quite often left in the racks rather than risk disturbing working circuits by attempting removal.

Several steps have been taken to relieve this problem: new and smaller cables have been developed for use in the system, a group distributing frame has been developed to reduce recabling requirements, the equipment arrangements have been designed to combine many functions in one bay thus reducing interbay cabling, the floor plan layouts have been changed to provide arrangements which will keep interbay cabling runs within reasonable lengths.

7.1 Cables

Previously a large number of cables were required between the various bays in an L multiplex terminal. Not only were these large in number but also relatively large in size. The shielded pair cables connecting the channel banks to groups have been reduced to one-half the diameter by the use of the new 761A cable rather than 720 cable, as shown on Fig. 39. A significant reduction in size was achieved by the use of miniature coaxial cable, 0.1 inch in diameter. The miniature cable is one-fourth the diameter of the 724 cable previously used. With the smaller size comes a much smaller bending radius which has made it possible to

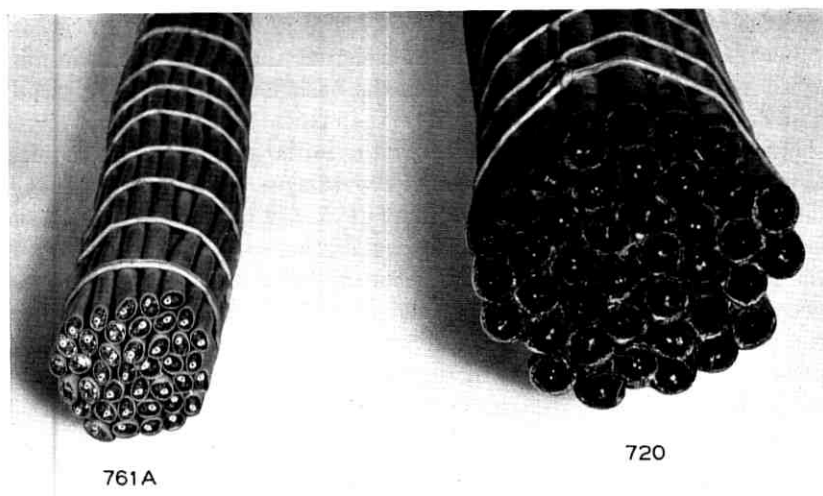


Fig. 39 — Size comparison of 40 shielded pairs of 761A vs 720 cable.

place this cable into tight areas. However, the transmission loss of the miniature coaxial cable limits its use to runs of less than 100 feet.

7.2 Bay Layouts

The earlier terminals required one bay for each supergroup bank and one bay for each two group banks. The group carrier supply required one bay to supply each 10 group banks. In conjunction with the active equipment units an additional bay containing jacks was required to provide a means for testing and emergency patching. This bay was located in a central test area, and all the input and output leads of the equipment were cabled to this location.

In planning the new design it became apparent that large savings in cabling could be realized by combining many functions in one bay. With the use of transistors, ferrites, and other new devices, the necessary reduction in size of the apparatus and equipment units has been realized.

One bay, 11 feet 6 inches high, 19 inches wide, is required for each direction of transmission for the group and supergroup equipment for 600 channels. All of the needed carrier supplies are also accommodated on the same bays. In the earlier designs this much equipment required more than 20 bays.

In addition to providing the standard features, these same two bays contain all of the automatic gain regulators and the testing, maintenance, and alarm circuitry, most of which was not available in the earlier equipment. The transmitting bay for the L600 multiplex is shown on Fig. 40 and the receiving bay on Fig. 41.

One of the major considerations in designing the bays was to minimize options and include as many units in the basic bay as was economically and functionally feasible. This means that units such as the group bank shelves are usually provided whether the initial circuit layout requires all of them or not. This technique of providing a complete bay package makes it possible to have the greatest flexibility of circuit changes and additions with the shortest possible installation intervals. By providing the framework and hardware for an entire 600 channels in the bay the necessary plug-in units can be added as growth dictates. The cost for carrying the initially unused shelves can be justified by the savings possible by having the units factory-wired and the savings in time and convenience in having the equipment available when needed. The plug-in equipment is omitted until needed.

7.3 Plug-in Unit and Bay Hardware

The active circuits, which affect transmission, are contained in modular plug-in units mounted in the same shelves with the patching jacks.

The necessary gain adjustments are contained in these plug-in units, are accessible from the front of the bay, and can be adjusted on an in-service basis. The active spare transmission units, scanner, carrier supply, carrier distribution equipment, and alarm units mount in the bay or bays of transmission equipment they serve.

7.3.1 *Receiving Modules*

The group and supergroup receiving modules each consist of amplifiers, networks, filters, and miscellaneous apparatus assembled in a drawer-like metal chassis of the plug-in type. The physical appearance of the two units is quite similar except for the front face plate. The group module with one side cover removed is shown on Fig. 42. This is the reverse side from that shown on Fig. 16. The metal chassis, which houses the apparatus, consists of an I-shaped aluminum extrusion for the basic shell, a die cast front plate, a rear plate perforated for ventilation and two removable side plates. The apparatus units contained in the shell are mounted to the web of the extrusion by screws. The units are then interconnected by soldered wires. Any unit, however, can be removed independently by disconnecting the wires and removing the mounting screws.

The front plate contains jacks for testing, a potentiometer for adjusting the regulated output power and an alarm lamp which indicates loss of pilot. The connector plug is attached to the front plate and is cabled up through the shell to the individual apparatus units.

The module is locked in place on the equipment shelf by means of a captive screw located on the face plate and a tang located at the rear of the shelf which engages the rear plate. No other guides are required on the chassis or shelf for alignment purposes. This entire package, with all its circuitry, is only 5 inches high, by $2\frac{3}{4}$ wide and 15 inches long. Each shelf contains one supergroup module and the associated 5 group modules. The receiving bay contains 11 shelves, one of which holds the spare group bank equipment.

7.3.2 *Transmitting Modules*

The transmitting group assemblies consist of an amplifier and a modulator housed in a single apparatus can which connects to the equipment shelf through a multiple-contact plug. Five group amplifiers and one intermediate amplifier are mounted on one shelf as shown on Fig. 43. The amplifiers can be extracted by using a special tool which inserts in the slot on the front face. The associated group bandpass filters are mounted in the rear.

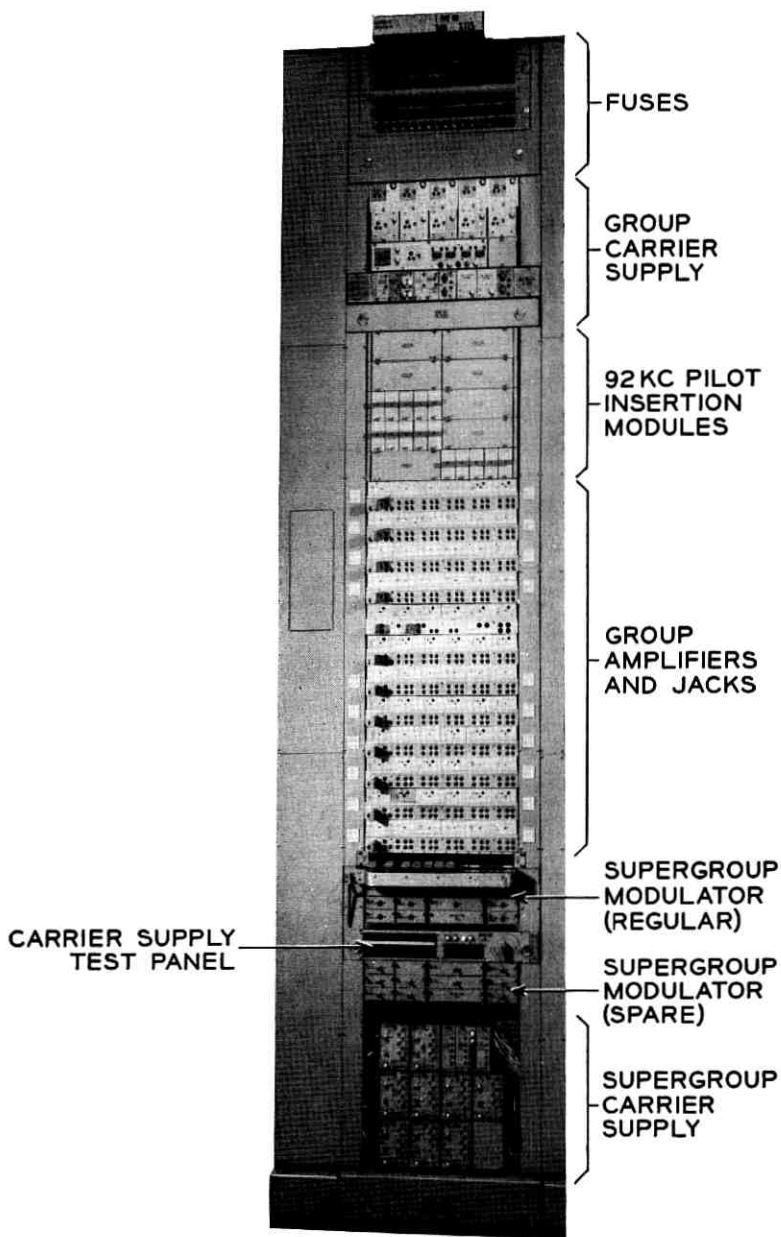


Fig. 40 — Transmitting bay, L600 multiplex.

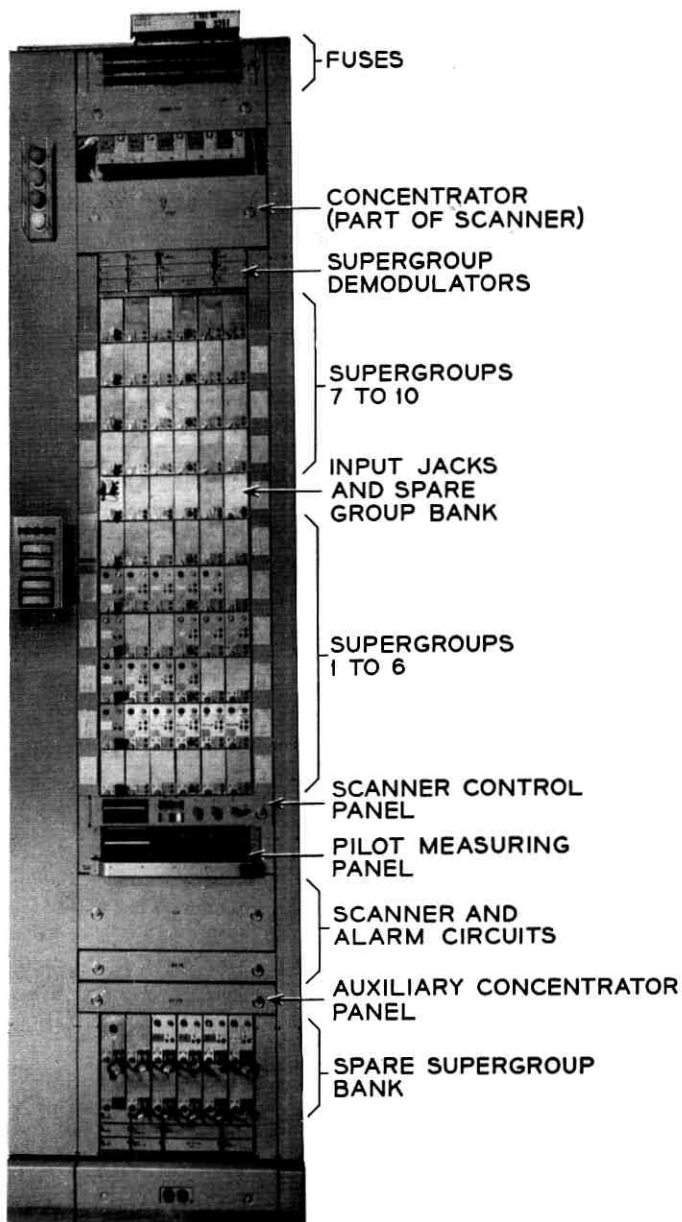


Fig. 41 — Receiving bay, L600 multiplex.

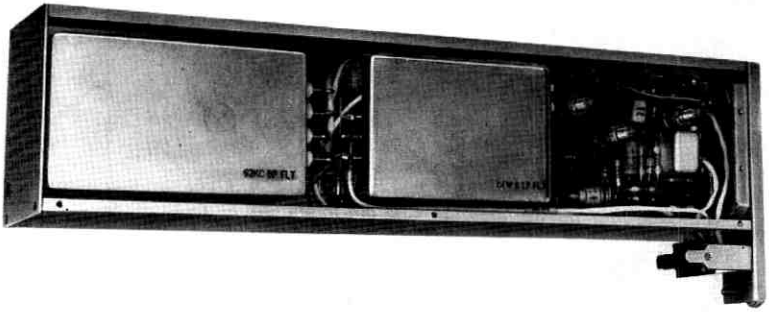


Fig. 42 — Receiving group module with cover removed (reverse side from that shown on Fig. 16). Differential dc amplifier and alarm circuit on right.

The supergroup modulator, demodulator, and bandpass filter units are mounted on flat plates which slide into a shelf framework shown on Fig. 44. The same basic design is used for both the transmitting and receiving bays. Interconnections between these units are made by the use of miniature coaxial cable equipped with jacks.

7.3.3 Other Equipment

The remaining equipment units located in the bays are primarily flat-panel construction packaged to utilize the 15 inches depth. These units are the fuse, concentrator, pilot measuring circuit, scanner and control panel, the alarm and auxiliary concentrator panels.

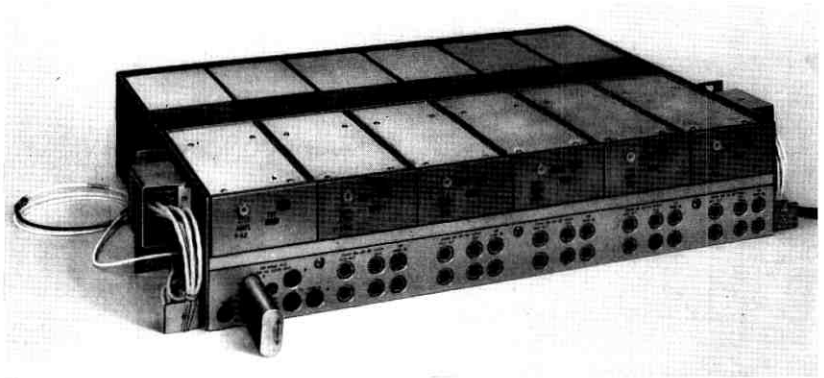


Fig. 43 — Transmitting group bank. Plug-in unit on left is group intermediate amplifier. Others are 5 group amplifiers and modulators. Group bandpass filters are in rear.



Fig. 44 — Supergroup modulators and bandpass filters.

7.4 Group Distribution Frame

With the continuous growth in the toll plant it frequently becomes desirable to transfer groups of circuits from one system to another. For example, when a new microwave or cable system is completed, some circuits are transferred to it from existing systems. This provides capacity for growth on the existing systems as well as on the new system.

Most toll offices now have a main distribution frame where all individual voice circuits in the office are connected to the voice frequency patch bays and through them to the channel banks. The individual circuits can be transferred by changing the connections on this frame. When whole groups of twelve voice channels were to be transferred it has been the practice to make the transfer by changing the cables between the channel banks and the group equipment, rather than transferring twelve pairs in the main distribution frame. When the reassignments involved group connectors, no alternative existed except to recable the connection from the connectors to the group equipment.

A study of the history of circuit reassignments indicated that each channel bank or group connector was reassigned on the average of once in three years although there are wide variations. Several possible methods of providing a convenient means for making these transfers were studied. All involved running all channel bank and group connector cables to a common location in the office and fanning out from there to all the group equipment. At the common location, some form of cross-connection must be provided.

Providing distribution frames for the group connections permits installation of channel banks and group connector filters without pre-

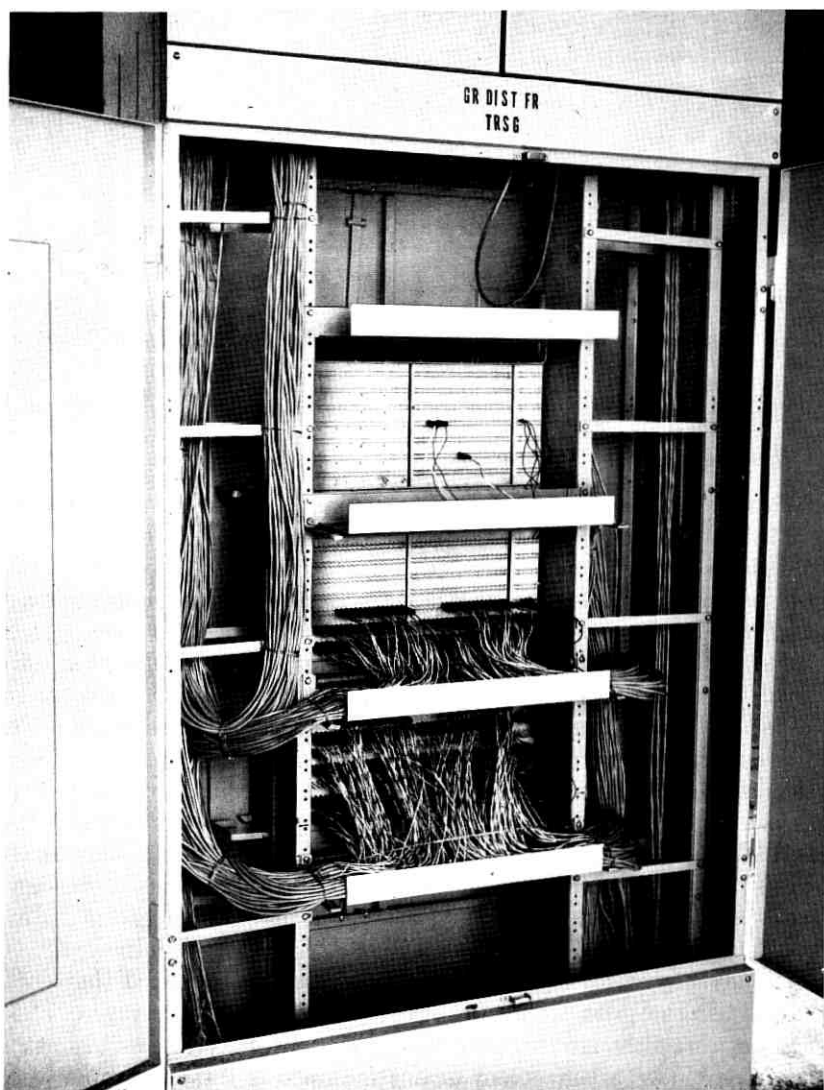


Fig. 45 — Group distributing frame for cross-connecting up to 1200 groups, cabling side.

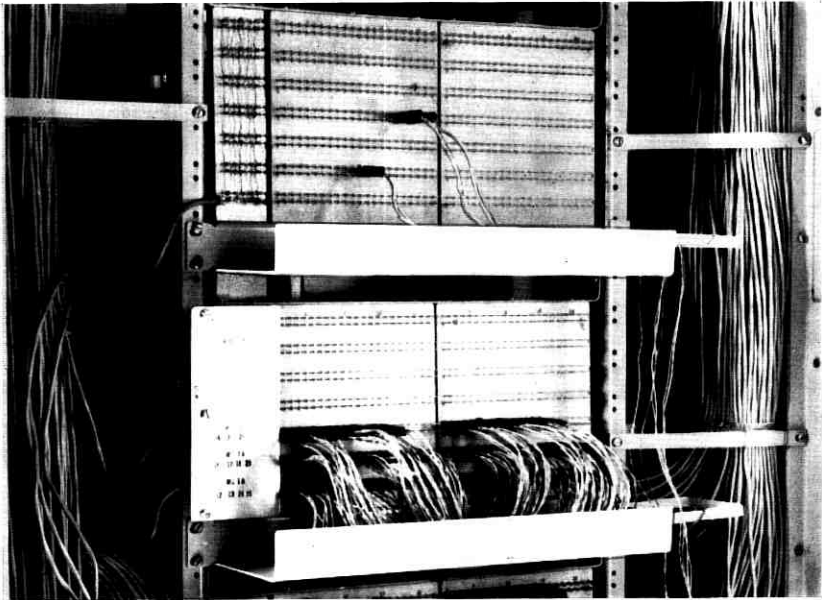


Fig. 46 — Group distributing frame, cross-connection side.

determining to what group equipment each will be connected, simplifying engineering and installation planning.

Tests of crosstalk indicated that unshielded pairs could be used for the cross-connections provided the whole frame was shielded from external disturbances. Wire-wrapped connections to terminals were used instead of pin-type connectors to conserve space and shorten the length of the cross-connections.

The frame shown on Fig. 45 can accommodate up to 1200 channel bank connections for one direction of transmission. The other direction is accommodated on a second bay. The cabling is run on both sides of the terminal strips and terminates on the reverse side of the terminal strips. All cross-connections are made on the front of the terminal strips. A close-up of one terminal strip is shown on Fig. 46. The insulated tubing over each pair facilitates locating a desired pair and reduces the risk of interfering with adjacent pairs when changes are made. Wire-wrapping and unwrapping tools are used.

The installation of group distribution frames involves recabling every channel bank or group connector in the office, but once this is done, further reassignment can be made with a minimum of effort.

7.5 Office Layout

To capitalize on the features discussed, new floor plan layout arrangements are desirable. These layouts are designed to keep cabling

runs to a minimum and combine certain functions by grouping the bays. The grouping is significant especially when certain units are shared by two or more bays. The scanner and pilot measuring equipment, spare supergroup equipment, and alarm equipment will serve up to three receiving bays while the carrier supplies will serve three transmitting bays and three receiving bays. The method of locating the bays in this manner will keep the cable runs, between the shared equipment, within the 100-foot loss limitation of the miniature coaxial cable. These bays are designed for front-side maintenance, making it possible to mount them back-to-back or next to a wall.

VIII. CONCLUSION

New group and supergroup terminals for long-haul telephone circuits have been developed which incorporate many new features to improve performance. The use of transistors and associated modern components, together with new materials, has permitted a great reduction in size and the promise of increased reliability. The equipment has been designed to reduce installation effort and minimize office cabling. The addition of group and supergroup regulation in the receiving terminals will make possible new high standards of transmission stability.

The improved performance and lower installed cost of the new terminals should assure a wide use for several years.

IX. ACKNOWLEDGMENTS

As in any project of this size, many individuals have contributed. In addition to the members of the authors' groups and the authors of the companion papers, significant contributions have been made by D. S. Williams, J. L. Garrison, D. W. Grant, S. G. Hale, and members of their groups.

REFERENCES

1. Hallenbeck, F. J. and Mahoney, J. J., B.S.T.J., this issue, p. 207.
2. Blecher, F. H. and Hallenbeck, F. J., B.S.T.J., **41**, Jan., 1962, p. 321.
3. Crane, R. E., Dixon, J. T., and Huber, G. H., Trans. A.I.E.E., **66**, 1947, p. 1451.
4. Albert, W. G., Evans, J. B., Jr., Ginty, J. J., and Harley, J. B., B.S.T.J., this issue, p. 279.
5. Elmendorf, C. H., Ehrbar, R. D., Klie, R. H., and Grossman, A. J., B.S.T.J., **32**, July, 1953, p. 781.
6. Bode, H. W., *Network Analysis and Feedback Amplifier Design*, New York, D. Van Nostrand Co., 1945.
7. Oliver, B. M., Proc. I.R.E., **36**, 1948, p. 466.
8. Kinzer, J. P., Trans. A.I.E.E., **68**, 1949, p. 1181.
9. Ketchledge, R. W. and Finch, T. R., B.S.T.J., **32**, July, 1953, p. 833.
10. Clark, O. P., I.R.E. Convention Record, 1962.
11. Caruthers, R. S., Trans. A.I.E.E., **58**, 1939, p. 253.
12. Aikens, A. J., and Lewinski, D., B.S.T.J., **39**, July, 1960, p. 879.

Carrier Supplies for L-Type Multiplex

By W. G. ALBERT, J. B. EVANS, JR.,
J. J. GINTY and J. B. HARLEY

(Manuscript received October 24, 1962)

This article describes the carrier and pilot supply equipment of the L-type multiplex. Also discussed is the modernized carrier supply for the A-type channel banks described in an earlier paper.

A short review of previous carrier equipment and recent developments in carrier application is followed by a general discussion of the design approach. Finally, specific circuits and equipment units are described.

I. INTRODUCTION

The L-type multiplex operates by frequency-division and thus requires a source of carrier power for each channel or group of channels transposed from one frequency band to another. Specifically, this means a carrier source for each channel, group, and supergroup modulator. These carrier frequencies must be resupplied at the receiving terminal also; and since the system is single-sideband with carrier suppressed, they must be resupplied with extreme frequency accuracy to prevent undue frequency shift in the channels. This is accomplished by accurately synchronizing these carriers to a received pilot which is derived from the carrier supply at the transmitting end.

II. REVIEW

2.1 *Old Carrier Supply*

Since a synchronized carrier supply tends to be expensive, it has been designed conventionally as centralized equipment so as to spread the cost over a great many channels. This was appropriate when the emphasis was placed on large installations. The old equipment (hereafter called "electron tube" equipment) was designed before large-scale application of carrier techniques to short-haul trunks. The new equipment recognizes this use of carrier and is designed to be economical for small-density installations without unduly penalizing the large.

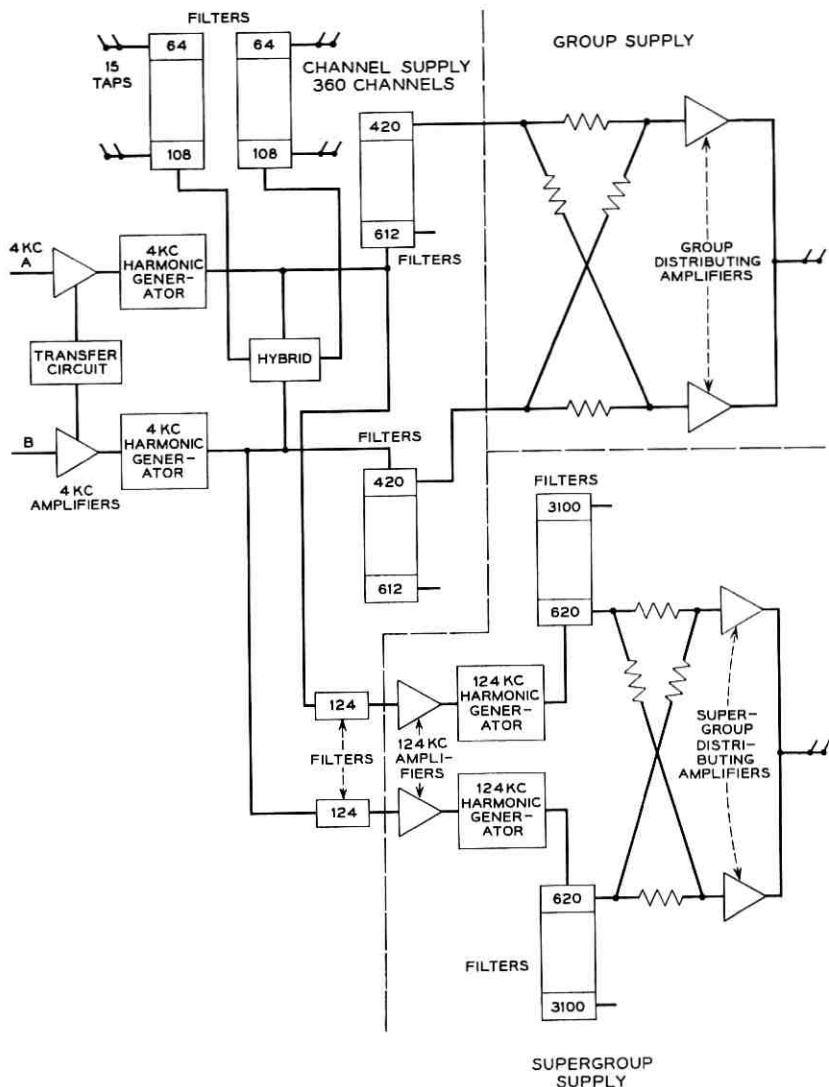


Fig. 1 — Electron tube carrier supply.

Fig. 1 is a simplified schematic diagram of an electron tube carrier supply. All carrier frequencies are multiples of 4 kc, the frequency spacing of the channels. These carrier frequencies are produced by driving two harmonic generators with separate 4-kc sine wave sources designated A and B. Only one of the harmonic generators is permitted to

operate at any time, the other one being disabled by the transfer circuit. The working harmonic generator feeds two similar banks of channel carrier filters via a hybrid transformer. Each pair of filters selects a carrier frequency and feeds it to as many as 30 channel modems. The figure is simplified, in that the even and odd harmonics actually come out of the harmonic generator on two different busses, and the filters are divided into even and odd sets, each fed by a hybrid.

The group carriers are all odd multiples of 4 kc, and are obtained in a manner similar to the channel carriers, except that duplicate filters are used rather than a hybrid, and distribution amplifiers are needed because of the power levels required by the group modulators.

For the supergroup carriers, the 124-kc harmonic component is filtered and fed to another harmonic generator which produces the supergroup carriers as multiples of 124 kc. Although not shown in Fig. 1, the equipment also generates 92-kc group pilots and 64-kc line pilots by the same techniques discussed above.

This electron tube carrier supply represents a highly centralized design. The interdependency of the "channel" carrier supply on the one hand, and "group" or "supergroup" carrier supplies on the other, is at once apparent; the latter will not operate without the former. Also, since the group and supergroup supplies serve up to 3000 channels, they are made extremely reliable by built-in redundancy. This redundancy is reflected all the way back to the 4-kc harmonic generators, which are duplicated and protected by the automatic transfer circuit.

The resulting protection of the 4-kc harmonic generators is also available to the channel carrier supply, even though it is not as necessary for its 360 channels as it is for the 3000 previously mentioned. This bonus protection comes at a price: each time a 360-channel mark is passed in the office-growth pattern, a new pair of harmonic generators complete with automatic transfer must be installed.

2.2 *Recent Developments in Carrier Application*

The electron tube equipment just described has been usefully and economically applied on a large-scale basis, mainly in multiplexing for 1860-channel L3 coaxial systems¹ and for 600-channel TD-2² microwave radio channels.* In the meantime, however, frequency-division techniques have been applied successfully to shorter and shorter trunks,

* The distinction should be emphasized between a voice or telephone channel, which is roughly 4 kilocycles in bandwidth, and a radio channel, which is many megacycles in bandwidth and carries many telephone channels.

and to smaller and smaller installations. A single N carrier³ 12-channel terminal in an office, for example, is entirely practical on an economic basis.

Similarly, the "ON" carrier terminal has been applied as a 96-channel multiplex for the TJ and TL short-haul microwave radio systems.⁴ The channel capacity is limited due to the loading effect of the transmitted carriers. The L-type multiplex, however, can be used to stack 240 or more channels for transmission on these radio systems, since carriers are not transmitted. The cost of the electron tube L-type equipment usually makes this prohibitive, unless the L carrier supply is already in the office and unused taps are available. The nature of the short-haul radio route tends to make this favorable circumstance improbable.

Thus, we see that the L terminal approach (single-sideband) has been technically but not economically attractive for the low density application of frequency-division techniques. Yet the low density carrier-derived trunk is today very common, and is becoming more so all the time. This fact has been paramount in setting the detailed design objectives of the new multiplex carrier supply (hereafter referred to as "transistor carrier supply"), and in addition has led to the development of the L60A and L120A packages.*

III. OBJECTIVES

The technical performance objectives of the transistor carrier supply, like those of the electron tube type, are set by the nature of the modulators it supplies. A carrier power considerably higher than the signal power is generally required by the modulators to enhance their linearity and to stabilize them against variations in carrier power. Thus, the essential requirements for the carrier supply are frequency accuracy and sufficient power, with a fairly lenient requirement on amplitude stability.

In the case of carrier supplies for channel modulators, these requirements are somewhat modified because of an additional objective: the channel modulator should have a specific amplitude limiting characteristic. To allow limiting at the proper signal level, the channel modulator is supplied with carrier power considerably reduced with respect to that supplied a group or supergroup modulator. Since the channel

* This paper describes the L600A and L1860A carrier supplies, as well as channel carrier supplies. The L60A and L120A circuits are similar, but the channel carrier supply is integrated with group and supergroup supplies in one package.⁶

modulator is thus made more sensitive to carrier level variations, the channel carrier supply must be more stable. This is more fully described in Section 5.2.

The essential requirements discussed above do not differ from those satisfied by the electron tube carrier supply. However, additional ones were established: lower cost, easier maintenance, miniaturization, and general modernization. Finally, a very specific objective formed early in the design was decentralization of the carrier supply, and integration with the equipment it serves. This represents an important step in making the new multiplex attractive for small installations, and is also consistent with the general desire to reduce the volume of interbay cabling. The decentralization objective must, of course, be tempered with the continuing need for economy in large density installations.

IV. NEW MULTIPLEX CARRIER SUPPLY

4.1 *Design Approach*

In accordance with the decentralization objective, the transistor carrier supply consists of two basic equipments: the channel carrier supply, which is located with the 30 channel banks it serves (three bays); and the group and supergroup supply, which mounts in one of the six multiplex bays it serves.

It may be inferred from the above that the multiplex is still not completely integrated, since the channel banks and the rest of the multiplex tend to be independent. This is as it should be, since most offices do not terminate all channels at voice-frequency, but pass some through the office at group or supergroup frequencies. In the case of the transistor multiplex, it is no longer necessary to install parts of a channel carrier supply merely to drive group and supergroup carrier supply equipment.

Although the group and supergroup supply is common to six bays (three transmitting and three receiving), each of these bays has its own local group carrier distributing amplifiers and busses. This reduces the number of cables from the group carrier supply proper from one per modulator to one per frequency, thus effecting a cable reduction of 50 to 5. Although it is primarily a cable-saving device, the local distribution approach also reduces the power requirement on individual transistor output stages by about 8 db by virtue of distributing a given carrier via six secondary amplifiers rather than one.

Channel bank bays not containing a carrier supply could benefit by

a 10 to 1 reduction of interbay cables if local distribution arrangements were provided for them. The arrangement chosen, however, (no local distribution busses) obviates the need for amplifiers completely, except for the one associated with the harmonic generator. This is possible because of the modest power requirement of a channel modem as compared with that of a group or a supergroup modulator, as mentioned before. Thus, the design is an economical and reliable one, and the required cables normally feed three adjacent bays, and so do not enter the cable racks.

4.2 *Reliability*

The question of carrier supply reliability is an interesting one, and must be discussed together with maintenance and transmission hits, these three subjects being intimately related.

The electron tube carrier supply, being common to as many as 3000 channels, was developed to satisfy an extreme reliability requirement. Consequently, it has much redundancy and automatic switching. For example, each 360 channels of carrier supply is duplicated and automatically protected, largely because it contains equipment for group and supergroup supplies for up to 3000 channels. In the new design, the carrier supply amplifiers are not automatically protected, or duplicated on a one-for-one basis, unless they serve more than 360 channels. This is because the group and supergroup carrier supplies are independent of the channel carrier supply, and also because the inherent reliability of the transistor amplifiers is expected to be substantially better than that of the electron tube amplifiers.

The new transistor amplifiers are plug-in, and are thus readily replaced in case of failure. Amplifiers serving 360 channels or less, though not duplicated, are protected by spares common to several bays. These spares are activated and provided with alarms so as to be readily available if needed.

Amplifiers common to more than 360 channels are automatically protected on a one-for-one basis. The switches used for this are not hitless, but are activated only upon a failure. They are not used for routine maintenance, since this is performed on working units on an in-service basis. Thus, the need for frequent switches has been removed, rather than incorporating an expensive hitless arrangement.

4.3 *Description*

The transistor multiplex equipment (L600A or L1860A) is used as all or part of the terminal equipment for several types of line facility. For

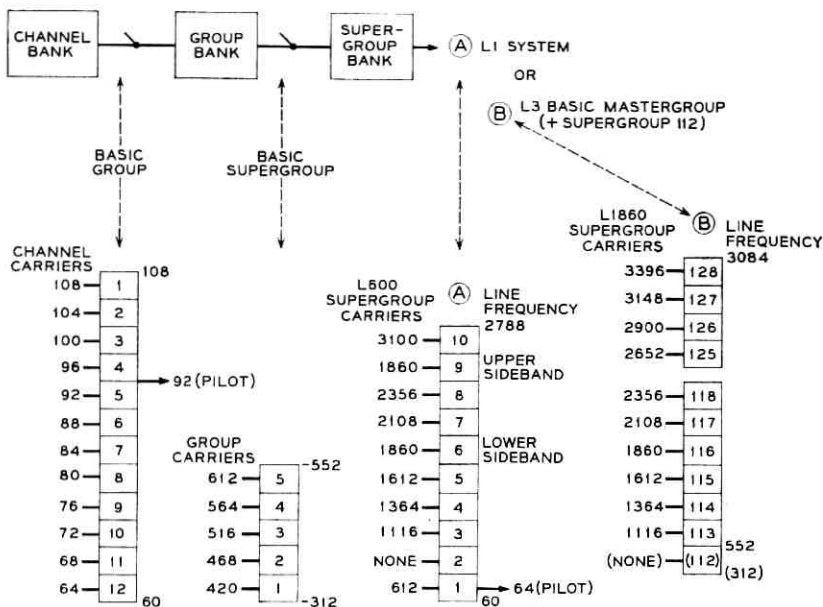


Fig. 2 — Line and carrier supply frequencies.

present purposes, these facilities may be considered to have either an L1 or an L3 line frequency assignment. The line frequencies and the required carrier frequencies are depicted in Fig. 2. Here an L1 system is assumed to include a 600-channel (ten-supergroup) facility—for example, a TD-2 radio channel. An L3 system is considered to be an L3 coaxial line or a TH radio⁵ channel, 1860 channels (31 supergroups) in either case. For the L3 system, an additional two steps of modulation are required to place the two upper mastergroups at line frequency.* These additional steps of modulation are provided by L3 mastergroup equipment which works with, but is not a part of, the transistor multiplex. The master group equipment is presently of the electron tube type.

The transistor multiplex for the L1 system is called “L600A multiplex,” while that for the L3 system is called “L1860A multiplex.”⁶ Fig. 3 shows the L600A multiplex carrier supply. The channel carrier supply makes use of harmonics of 4 kc as in the old supply, with differences in the protection features and isolation from the other carrier supplies as previously discussed. In addition, a single set of 12 filters† is used for all 360 channels, thereby reducing the number of filters required by a factor of two.

* An L3 mastergroup consists of ten or eleven supergroups.

† Carrier supply filters are described in a companion paper.⁷

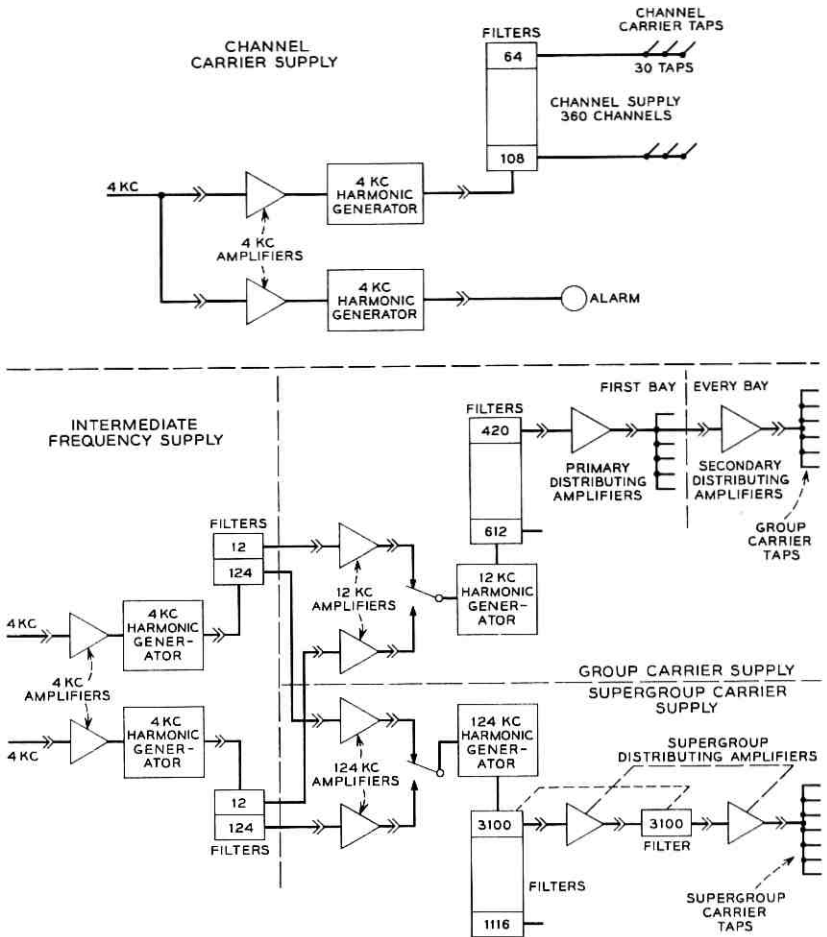


Fig. 3 — L600A carrier supply.

The balance of the carrier supply has its own pair of 4-kc harmonic generators, fed from additional taps on the 4-kc primary frequency distributing bus.* These generators are used to derive only four frequencies: 12, 124, and (not shown) 64 and 92 kc. The former two frequencies form basic frequencies for the group and supergroup carrier supplies, respectively, while the 64 and 92 kc represent line pilot (as

* A new 4-kc primary frequency supply has also been designed, and is described in a companion paper.⁸

required) and group pilot respectively. The equipment unit supplying the above four frequencies is called the "intermediate frequency supply."^{*}

The 12-kc filters and amplifiers are duplicated. A relay selects either output to feed the harmonic generator (saturable inductor), and automatically switches to the other output in case of a failure. It should be noted that the group carrier frequencies are derived as harmonics of 12 kc rather than 4 kc. This results in easing the discrimination requirements on the group carrier filters, with attendant filter economies. The group carrier filters feed primary distributing amplifiers, which distribute the five frequencies to six bays. From there, the local (secondary) distributing amplifiers feed the carriers to all group modulators in the particular bay.

The operation of the supergroup carrier supply is generally similar, except for detailed differences in the distribution amplifier arrangements, the different base frequency (124 kc), and the different mode of generation of the upper four carriers in the L1860A application, these carriers not being multiples of 124 kc.

V. CIRCUIT DESIGN

5.1 *Introduction*

The individual carrier supply circuits consist of harmonic generators, filters, single-frequency drive amplifiers, and wideband distributing amplifiers. The passband characteristics of the latter are such that a single group (or supergroup) amplifier will handle any group (or supergroup) frequency.

Since the same basic saturable-inductor harmonic generator is used in several parts of the carrier supply, its operation is described only in connection with the group carrier supply.

The channel carrier supply description includes a discussion of the translation of system objectives to a requirement on drive amplifier stability. This setting of requirements, although obviously necessary for all circuits, is given only once as an example.

5.2 *Channel Carrier Supply*

The channel carrier frequencies, harmonics of 4 kc from 64 to 108 kc, are derived from the channel harmonic generator circuit by means of an

^{*} As described later, the intermediate frequency supply for the L1860A provides an additional frequency of 80 kc.

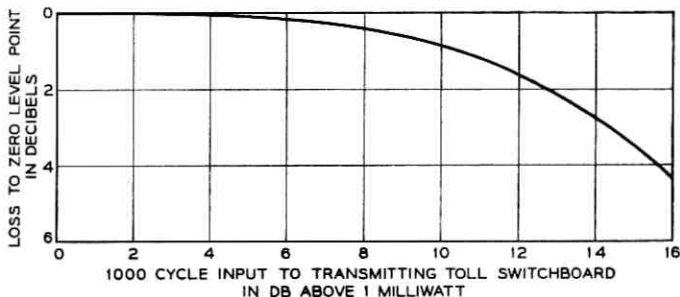


Fig. 4 — Limiting characteristic of channel modulator.

amplifier, pulse-forming network, and filter circuit. Distribution of the carriers is accomplished through a 30-tap capacitive distribution circuit.

Amplitude stability of the channel carrier supply is of prime importance if the system is to meet the transmission requirements of Direct Distance Dialing and future data systems. For most modulators, it is customary to supply carrier power at a much higher level than the signal power. This has two beneficial effects:

1. Variations in carrier amplitude with typical modulators are reflected as loss variations in the transmission path by only about one part in ten; i.e., a 1-db change in carrier power results in only 0.1-db change in transmission. This effect is called a "stiffness of 10:1."
2. High order ($c \pm mv$) products which frequently overlap to produce interchannel crosstalk are reduced.

The channel modulator, however, performs one additional function which changes the requirements and calls for much lower carrier-to-signal power. This additional function is power limiting, and is discussed below.

In single-sideband suppressed carrier systems, the system load depends on the number and power of the talkers active at any one time. Provision must be made to limit automatically the power of each talker so that overload of the system will not occur if, during the busy hour, an inordinate number of loud talkers is present at one time. Power limiting also protects the system from test tones accidentally applied at too high a level.

To achieve the necessary limiting, the carrier power provided to the channel modulators is 0 dbm* per modulator. This gives the desired limiting characteristic as shown in Fig. 4. Undesirable modulation products due to the low carrier power are reduced to acceptable levels

* db above one milliwatt.

by the channel bank filters, and the limiting effect on loud talkers does not reduce intelligibility or even impair naturalness to any appreciable extent. The low carrier power does, however, reduce the "stiffness" to a ratio of approximately 2:1 and thus imposes tight requirements on channel carrier stability.

The requirement on amplitude stability for the channel carrier amplifier may be derived in the following manner.

The 4000-mile requirement for message system net loss stability in order to satisfy the needs of Direct Distance Dialing will be taken as:

1. Sudden step variations in level shall be less than ± 0.25 db.
2. Average drift variations (bias) of the losses of a group of trunks from design value shall be less than ± 0.25 db with a distribution grade (standard deviation) from this average of less than 0.75 db.

If one allocates one-half of the bias to the line and one-half to the terminals that make up the 4000-mile channels, the bias (X) of any one terminal will be given by

$$\sqrt{n} X = 0.125 \text{ db}$$

where n is the number of links in tandem, and random distribution of link biases is assumed. If it is further assumed that the maximum number of links (n) in tandem is five, then

$$X = 0.056 \text{ db.}$$

What appears to be a stringent requirement on the amplifier is relaxed due to two things: the modulator stiffness, though not great, is 2:1, and the harmonic generator has its own limiting action which produces an additional stiffness of 5:1. Thus the requirement derived above is relaxed by some 10:1, and each amplifier should therefore exhibit an amplitude stability of 0.56 db.

Fig. 5 shows the amplitude stability of the 4-kc drive amplifier vs variations in temperature and office battery voltage. The loss variation of a typical channel carrier filter vs temperature is of the order of 0.2 db over the expected temperature range (40 to 120 degrees fahrenheit).

Fig. 6 is a simplified circuit schematic of the amplifier. Q_1 is a 16A diffused-silicon transistor used in the common-emitter configuration. This is followed by Q_2 , a 20C transistor in a phase splitting circuit which drives Q_3 and Q_4 , two 16A transistors used as emitter followers. These drive the output stage consisting of Q_5 and Q_6 , two 20G transistors in a class B output circuit. The transformer L1 provides series feedback around the amplifier to stabilize the circuit for variations in the μ path.

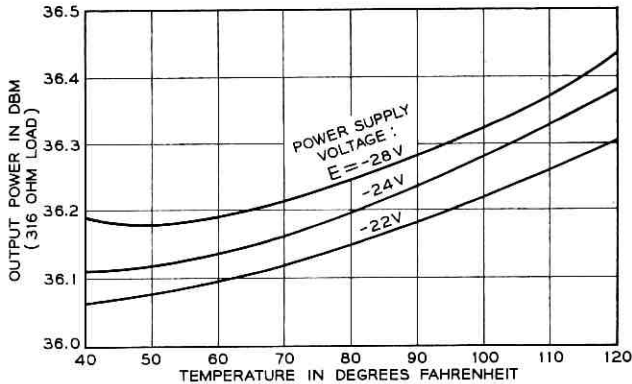


Fig. 5—4-kc drive amplifier—output power vs temperature and office battery variations.

5.3 Intermediate Frequency Supply

The intermediate frequency supply circuit is used to generate the basic frequencies required for the group, supergroup, and pilot supply circuits. The harmonic generators and filter circuits are duplicated and feed dual inputs to the group and supergroup circuits. Also included are the 64- and 92-kc pilot supply circuits. The pilot supply circuits contain amplitude stabilizers, since any change in transmitted pilot level would be interpreted (by regulators or maintenance forces) as a change in line or terminal gain. The stabilizers reduce pilot level variations of several db to a small fraction of a db.

The stabilizer is essentially a zero-gain circuit in which the pilot is amplified by two common-emitter stages, symmetrically clipped by diodes, and filtered to reject harmonics. The output power is held to within a few hundredths of a db for input variations up to ± 4 db.

The stabilizers for each pilot frequency are duplicated and switched. Resistive primary and capacitive secondary busses distribute the pilots to the appropriate circuits.

The L1860A intermediate frequency supply is the same as described above, except that 80 kc is produced as well as 124 kc for the supergroup drive frequencies. The 64 kc is not required, as the L3 line pilots are produced by L3 pilot supply equipment.

5.4 Group Carrier Supply

Five carrier frequencies are required to translate the basic group spectrum (60–108 kc) to the five assigned positions in the frequency

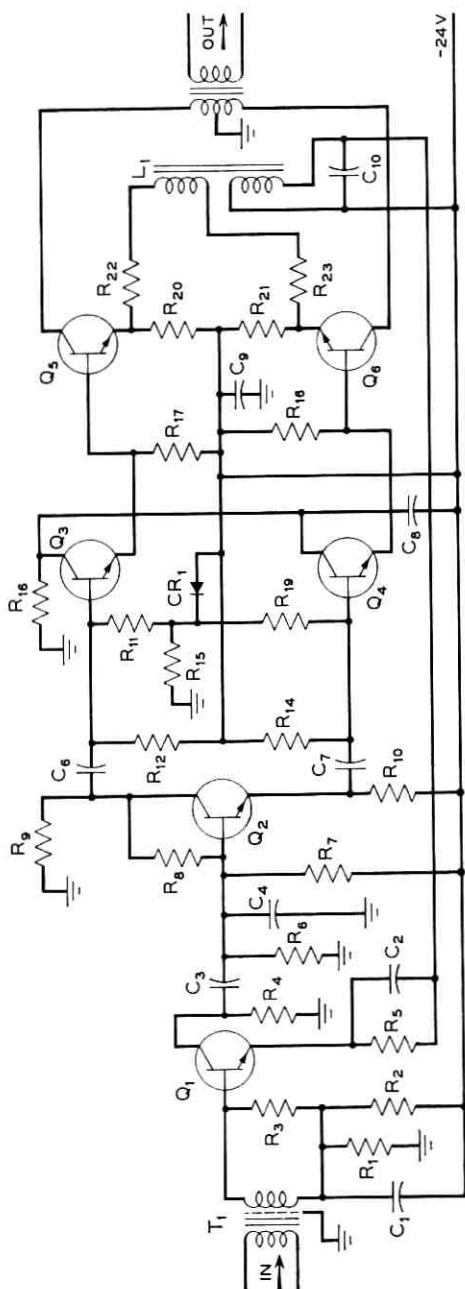


Fig. 6 — 4-ke drive amplifier schematic circuit.

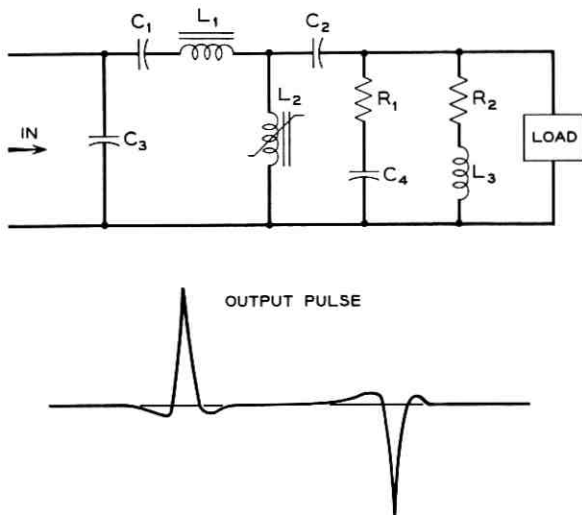


Fig. 7 — 12-kc harmonic generator circuit and waveform.

band of 312 to 552 kc; these carrier frequencies are 420, 468, 516, 564, and 612 kc, corresponding to groups one to five respectively. These frequencies are generated by the 12-kc harmonic generator circuit. The 12-kc amplifier which drives the harmonic generator is almost identical with the 4-kc amplifier previously described.

The harmonic generator circuit, shown in Fig. 7, consists of a high-Q series LC circuit C_1 , L_1 tuned to 12 kc, a saturable inductor L_2 , a differentiator circuit C_2 , and an impedance correcting network R_1 , C_4 , and R_2 , L_3 .

Pulses from the harmonic generator circuit are formed in the following manner. Initially, inductor L_2 is in a nonsaturated state and has a high impedance with respect to the series circuit L_1 , C_1 . This allows current to flow through L_1 , C_1 , and charge capacitor C_2 . When the charge is sufficiently large, inductor L_2 saturates, becoming a very low impedance with respect to the rest of the circuit. Capacitor C_2 rapidly discharges through L_2 and the load, producing a pulse of short duration* each half-cycle of the 12-kc drive frequency.

Fig. 8 shows the group primary distribution circuit consisting of a filter and pad circuit, a 232A amplifier, and a primary distribution circuit.

The 232A amplifier, shown in more detail in Fig. 9, is a two-stage

* Selection of pulse width is treated in Appendix I.

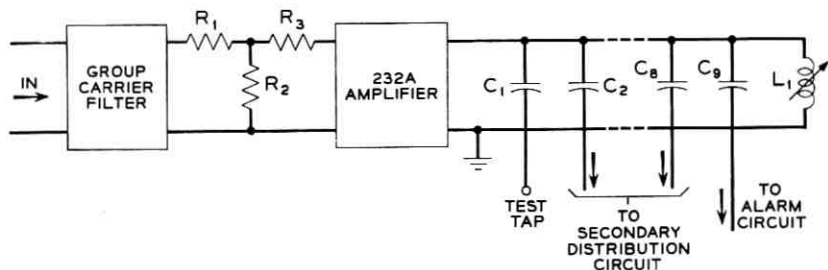


Fig. 8 — Group primary distribution circuit.

transistor circuit. The input stage utilizes a common-collector circuit while the output is a common-emitter stage. Local feedback, both shunt and series, is used in the output stage Q_2 . Potentiometer R_9 provides a gain control to vary the gain of the amplifier between 22.5 and 27.5 db. Voltage regulation diode CR_1 provides stabilization of the output stage against battery variations. Transformer T_3 connects the output of the amplifier to the primary distribution bus.

The primary distribution bus (Fig. 8) supplies the group carrier frequency to six working taps, one test tap, one alarm tap, and one spare tap. The six working taps supply six group secondary distributing circuits, one located in the same bay, the others in five other bays. Inductor L_1 is used to cancel the capacitive reactance of the tap capacitors and present a good resistive load to the 232A amplifier.

The group secondary distribution circuit, shown in Fig. 10, consists of a 230A amplifier and a distributing bus. It receives the group carrier

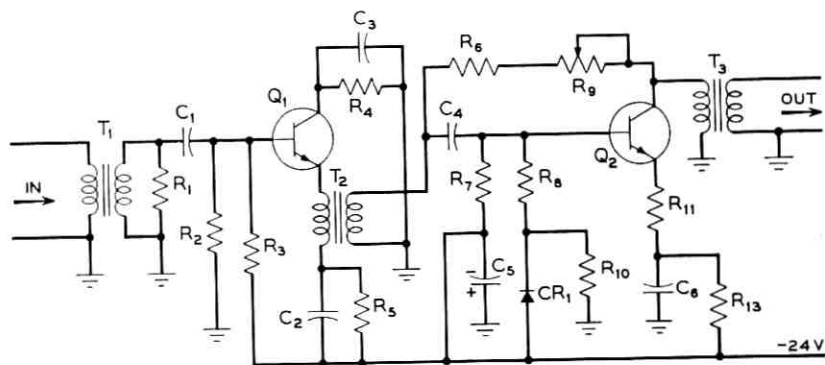


Fig. 9 — 232A amplifier circuit.

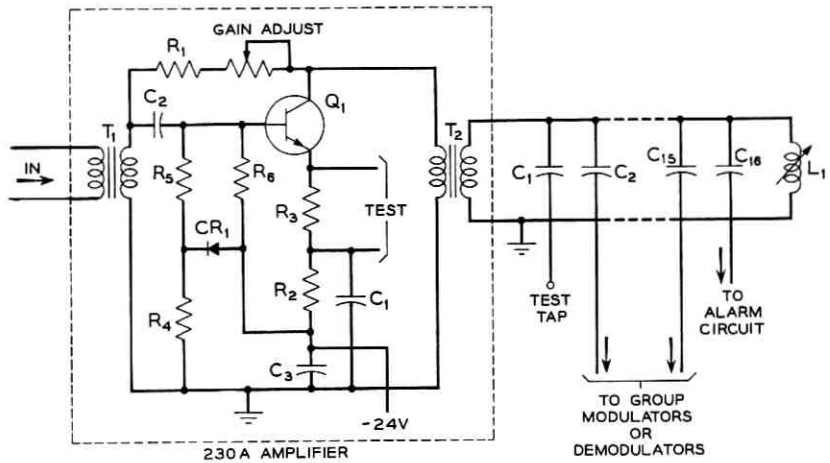


Fig. 10 — Group secondary distribution circuit.

frequency from the primary circuit and distributes it locally to the group modulators or demodulators as required. The amplifier is a single stage of power amplification required to raise the power to distribute locally in the bay. The circuit is a common-emitter stage employing both shunt and series feedback. Gain control is in the shunt feedback path to adjust the gain between 9.5 and 14.5 db. Stabilization of the gain of the amplifier against battery voltage variation is attained by use of a zener diode (CR_1).

The distribution bus has 16 taps: 13 working, one test, one alarm, and one spare. The group 5 spare tap feeds a Supergroup 1 modulator or demodulator.

5.5 Supergroup Carrier Supply

Like other parts of the carrier supply, the supergroup carrier supply is intended for use in both L600A and L1860A terminals. In the case of the supergroup supply, however, certain parts are optional, having application to a specific terminal.

When used with the L600A terminal, a 124-kc harmonic generator with appropriate filters makes available seven frequencies ranging from 1116 kc (SG3) to 3100 kc (SG10) as shown in Fig. 11. An eighth supergroup frequency, 612 kc (SG1), not a harmonic of 124 kc, is furnished by the group carrier supply. One supergroup requires no carrier (SG2), and two supergroups share one carrier frequency (SG6, SG9).

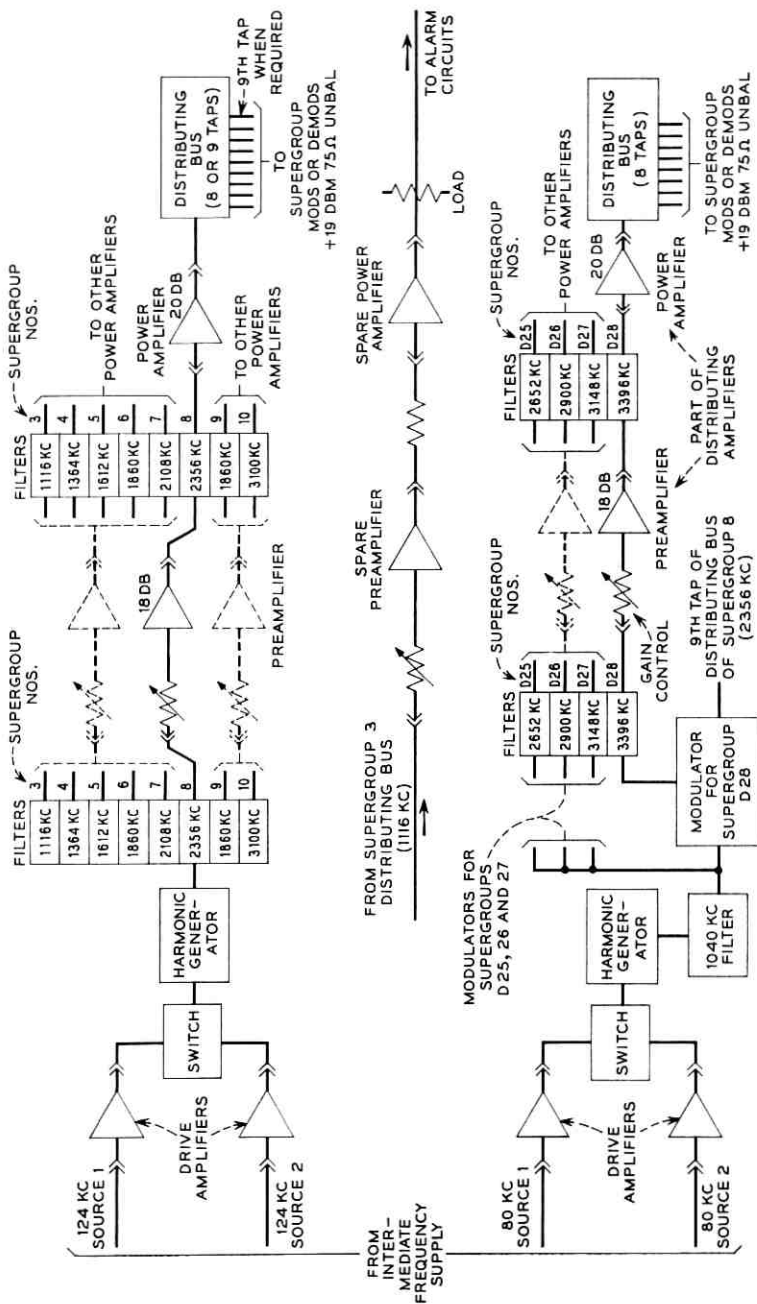


Fig. 11 — Supergroup carrier supply.

In an L1860A multiplex, four other frequencies shown in Fig. 11 are required which are not harmonics of 124 kc. These frequencies are obtained by mixing 1040 kc with the frequency used also for supergroup 5, 6, 7, or 8. An additional harmonic generator with an input frequency of 80 kc is made available as a source of 1040 kc. The 80-kc and 124-kc inputs required for the supergroup supply are provided by the intermediate frequency supply as previously described.

Both harmonic generators,* like the 4- and 12-kc generators, make use of saturable inductors to generate the pulses from which the supergroup frequencies are derived. Consider the 2356-kc frequency, for example. The small size of the saturable inductor limits the output of this frequency to +4 dbm at the output of the first section of the filter shown in Fig. 11. The harmonic generators are driven by an amplifier which will be referred to as the "drive" amplifier. The design of the drive amplifier is very similar to that of the power amplifier of the distribution module described later. The output of the amplifier is stepped up in voltage in the double-tuned circuit which couples the amplifier to the saturable inductor. The selectivity of this circuit causes the saturable inductor to be driven by a symmetrical wave, thereby equalizing the spacing between positive and negative pulses to insure predictable amplitude of the desired harmonics. Bandpass filters are used to select harmonics corresponding to the desired supergroup frequencies; these suppress adjacent harmonics by about 70 db.

Each distributing bus of the supergroup carrier supply drives eight modulators and/or demodulators at a level of +19 dbm. Even though supergroups 6 and 9 share the same carrier frequency (1860 kc) these supergroups are supplied from separate amplifiers and distribution busses. Uniform operation of the amplifiers, resulting in lower power transistor junction temperature, and uniform test procedures outweigh the economy to be gained by doubling the output of a single amplifier at 1860 kc.

A distribution amplifier is shown in Fig. 12. An operating power gain of 37 db is required to raise the +4 dbm level of harmonic output to +29 dbm while overcoming an average 12 db loss in the gain control, second half filter and pad. Amplifiers meeting these requirements are referred to as distribution amplifiers. These are designed for plug-in connection to simplify the stocking of spare amplifiers, and have sufficient bandwidth so that the single design suffices for all supergroups.

The distribution amplifier is subdivided into two separate amplifiers: a preamplifier and a power amplifier. This arrangement allows some

* These operate similarly to the 12-kc harmonic generator described under "group carrier supply."

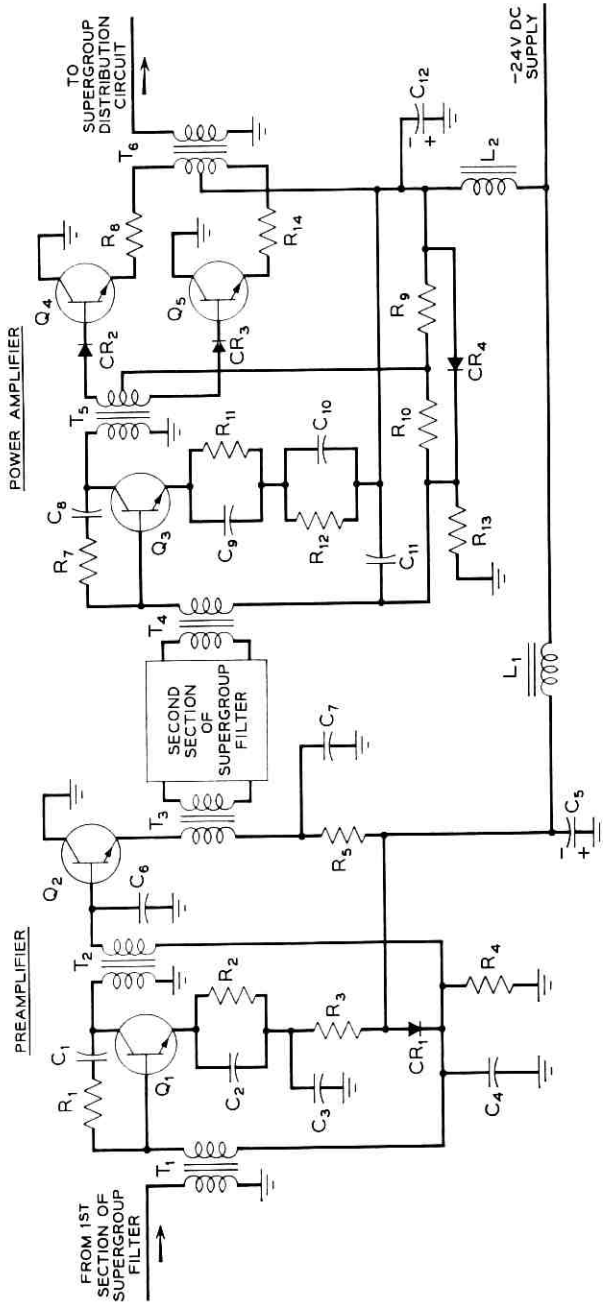


Fig. 12 — Supergroup carrier distribution amplifier circuit.

selectivity to be introduced between the two amplifier sections to filter out noise generated in the preamplifier and falling in the transmission band. To obtain this selectivity, the harmonic selection filter previously referred to is arranged as a two-part filter. The part preceding the preamplifier provides selectivity against harmonics, while the part following the preamplifier provides selectivity and also filters out noise.

The preamplifier is transformer-coupled with two transistor stages having an insertion power gain of 18 db. A frequency shaping network, part of the common emitter bias circuit of the amplifier, provides equalization over the supergroup carrier frequency range. The shunt (C_1 , R_1) and series (C_2 , R_2) feedback of the first stage are proportioned to control the impedance terminating the filter. The low output impedance of the common-collector second stage terminates the input side of the second half of the filter through transformer T_3 .

The power amplifier is a similar two-stage broadband amplifier with an operating power gain of 20 db. The first stage drives a class AB common-collector push-pull power stage. Under normal operating conditions the output power delivered by the amplifier is less than its power capability. By operating the push-pull output stage as a class AB stage a saving in emitter current is realized. Any reduction of current in the power transistors is important, as there is a corresponding reduction of transistor junction temperature. The junction temperature is kept within proper operating limits by attaching the common-collector power transistors through an aluminum block to the chassis frame.

The distribution circuit combines a tuned impedance transformation circuit with a capacitive type distribution bus. Tuning is provided to reduce harmonics. The distribution bus is designed to furnish +19 dbm to each of eight modulators and/or demodulators. Resistive loads are provided to terminate unused bus taps.

When the supergroup carrier supply is part of an L1860A multiplex, supergroups 9 and 10 are not required. Supergroups D25, D26, D27, and D28* are added and require, as previously mentioned, a source of 1040 kc for their formation.

The four added supergroup carriers (2652, 2900, 3148, and 3396 kc) are obtained by mixing 1040 kc with the carrier frequencies of Supergroups 5 through 8 (1612, 1860, 2108, and 2356 kc). The 80-kc harmonic generator required to generate 1040 kc differs from the 124-kc harmonic

* In the L3 carrier system, supergroups 25 through 28 were formed using two stages of modulation. In the application of L multiplex to L3 carrier systems, new carrier frequencies have been made available so these supergroups could be handled just as any other supergroup; for example, SG5 or SG8. The new supergroups are identified by the letter D, D25 through D28. New and old systems are compatible.

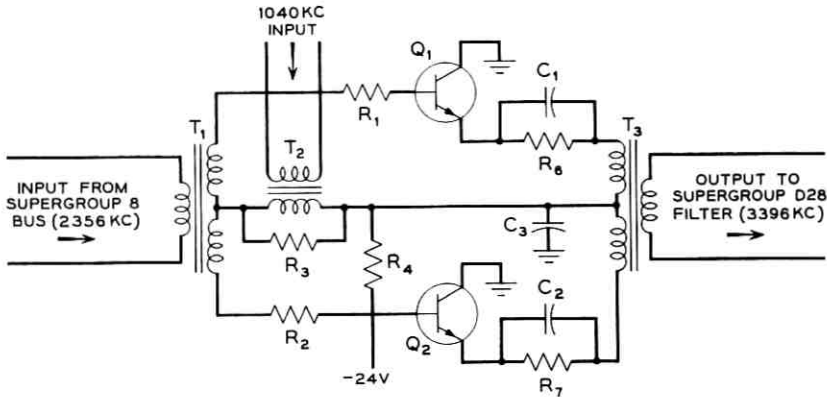


Fig. 13 — Modulators for directly formed supergroup carriers.

generator only with respect to frequency-selective components. The same drive amplifier may be plugged into either harmonic generator, and the protection and alarm circuits of both generators are identical.

The four modulators required to generate the new directly formed supergroup frequencies are identical; Fig. 13 shows the modulator for supergroup D28. A transistor modulator was chosen to secure some power gain, and this modulator is balanced to minimize interaction between 1040 kc and the supergroup frequencies with which it is mixed.

5.6 Protection and Alarm Features

A protection and alarm system is provided for the group and supergroup supplies, (1) to actuate office alarms whenever any carrier power falls below a preassigned level, (2) to light a lamp to identify the unit which has failed, (3) to provide automatic switching to a second source of input to the harmonic generator in cases of failure of either source, and (4) to provide a major alarm when all circuits associated with a harmonic generator fail.

With the exception of the channel carrier supply, which uses the 4-kc primary frequency directly, inputs to all L multiplex carrier supplies require frequency multiplication of the 4 kc in the intermediate supply. For maximum reliability, therefore, parallel circuits are provided. It would take a simultaneous failure of both amplifiers to disable the group and supergroup carrier and pilot supplies. Alarms are provided for each circuit to detect a failure.

The 12, 80, and 124-kc outputs of the intermediate circuit are each

double-fed through separate drive amplifiers and relay contacts to the group and supergroup harmonic generators. In each case the relay contacts select the output of only one amplifier to drive the harmonic generator. The relays are controlled by the output of the associated "drive" amplifier. The relay control circuit of the L600A supergroup supply, which is typical, is shown in Fig. 14. This figure has been simplified by the omission of contact protection and other circuit refinements not essential to the understanding of the circuit. It will be seen from this figure that a portion of the output of each drive amplifier is rectified for the operation of a relay. These relays (K_1 and K_2) are held operated as long as the outputs of their respective amplifiers do not fall below a preassigned level.

The output of the no. 1 drive amplifier is shown connected to the harmonic generator circuit through contacts of relay K_4 . Assume this drive amplifier fails; its associated relay K_1 releases, applying current to the polarized relays K_3 and K_4 . The contact springs, which move in the direction of the arrow when positive voltage is applied at (x), connect drive amplifier no. 2 to the harmonic generator. A second circuit closed by the release of K_1 lights the amber alarm lamp DS3, causing current to flow through pulse transformer T_1 . The build-up of current in the primary of T_1 causes a pulse of secondary voltage which passes through the diode gate CR₅ to the minor alarm circuit.

When the trouble has been cleared and drive amplifier no. 1 is reconnected, relay K_1 pulls up, extinguishing the amber minor alarm light. To avoid a service interruption at this time, relay K_3 , in the absence of current through its windings, holds its last operated position by means of a permanent magnet latch. It would take a current of the opposite polarity, as caused by the failure of drive amplifier no. 2, to cause the relay contacts to return to their original position. Amplifier no. 1 is not restored to service, but is connected to the green ready lamp DS₁, which doubles as a standby load. Amplifier no. 2 continues as the working amplifier until it fails.

Failure of both drive amplifiers, causing release of the K_1 and K_2 relays, permits current to flow through resistor R_3 and the winding of the major alarm relay K_5 . Operation of the major alarm relay lights the red major alarm light DS₅ and sends a pulse of current to the major alarm circuits through pulse transformer T_2 . It will be seen that the break contact of this relay supplies current to the minor alarm lamps of the supergroup distribution modules. These are locked out, since they would convey only misleading information.

The group and supergroup supplies each require an amplifier ahead of the distribution bus for each frequency. The output of each of these

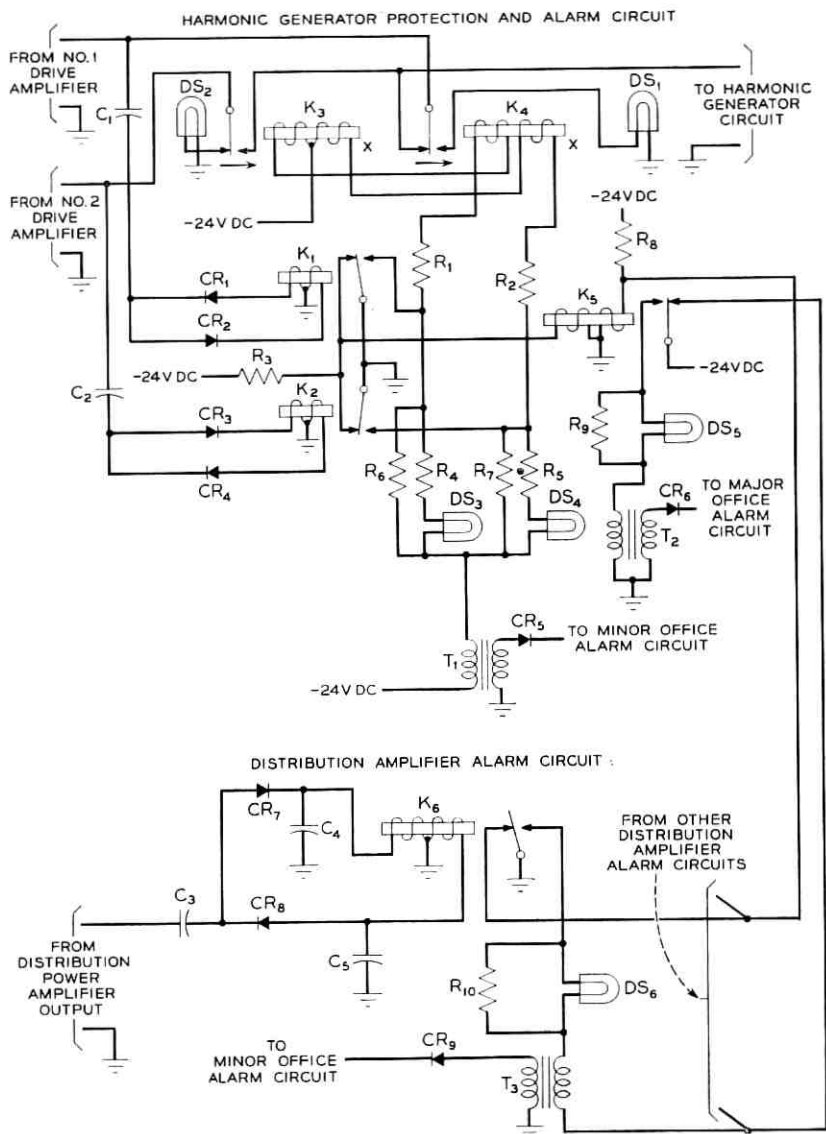


Fig. 14 — Supergroup carrier supply protection and alarm circuit.

distribution amplifiers is monitored by a full-wave rectifier and level-sensing relay similar to that already described. The supergroup distribution amplifier alarm circuit, which is typical, is also shown in Fig. 14. The make contact of relay K_6 of this circuit prevents operation of the major alarm relay K_5 . To operate the major alarm relay K_5 , all distri-

bution amplifier outputs must fail. When any distribution amplifier fails, the relay of its alarm circuit releases, lighting the amber minor alarm light DS_6 and causing current to flow in the primary of the pulse transformer T_3 . The secondary pulse passes through the diode gate to the minor alarm circuit. The major and minor alarm pulses referred to are converted to relay closures for operation of the standard office alarms by a centrally located office alarm amplifier. The principal advantage of the pulse alarm system is that each trouble sends but one pulse to the office alarm amplifier. In effect, every branch of the alarm circuit has an instantaneous automatic reset.

The control circuit for the 12-kc (group) and 80-kc (directly formed supergroup) supplies will not be described, as these control circuits are basically the same as those of the L1860A supergroup carrier supply.

Since the channel carrier supply is located with the channel banks rather than with the group and supergroup multiplex, the channel carrier alarms are arranged to couple independently to the standard office alarm system.

VI. EQUIPMENT DESIGN

6.1 *Introduction*

To complement the new carrier supply circuits, a completely new equipment approach has been conceived. This consists of an efficient modular packaging technique and, even more fundamentally, a decentralization of the equipment.

The old carrier supply design, because of the dictates of cost, component sizes, and reliability requirements pertaining to large offices, had been considered common equipment, analogous to the office battery supply. Separate bays were required for the pilot supply, channel, group, and supergroup carrier supply equipment. These bays were all dependent on the harmonic generator which was located in the channel carrier supply bay. Because of this and the stringent cable loss and crosstalk requirements, considerable constraint was placed on office layouts of L-terminal equipment. However, since floor space savings were not the primary objective of the old design, it was more than adequate. When we reflect on the large telephone office expansion of the past decade as well as the predictions of further increases, we realize that a reduction in the size of the carrier supply equipment and interbay cabling is now necessary.

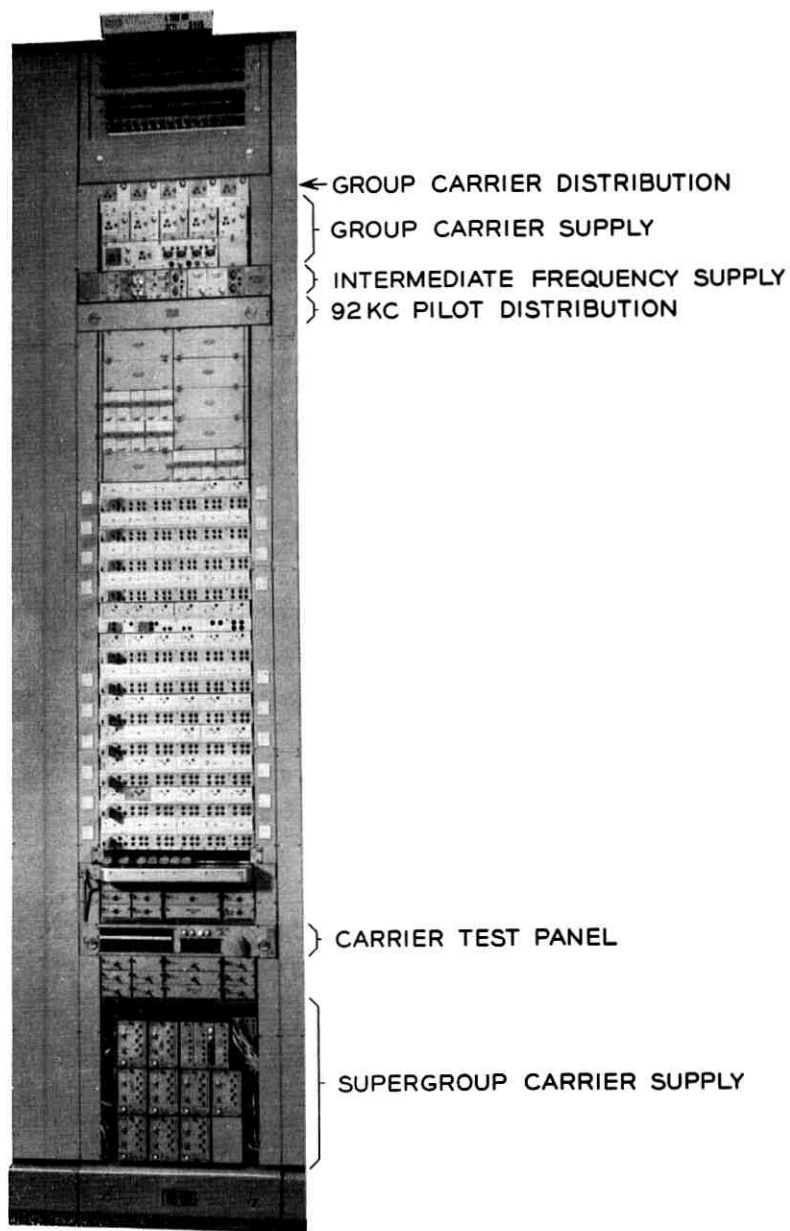


Fig. 15 — Carrier supply in transmit bay.

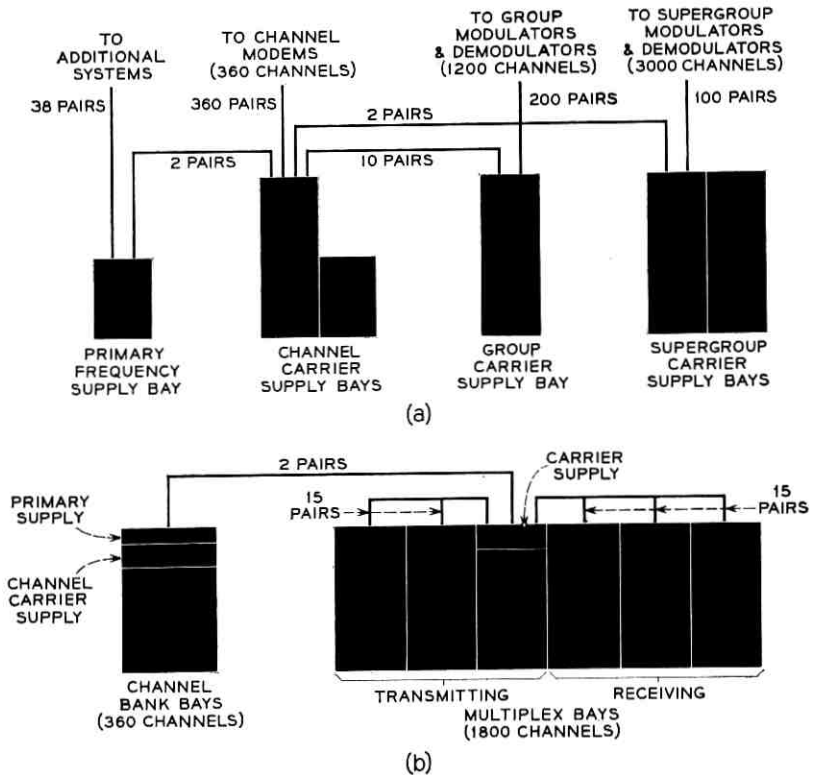


Fig. 16 — Cable comparison.

6.2 Bay Arrangements

The philosophy of decentralization, as applied to the transistor carrier supply designs, consists primarily of combining the carrier and pilot generators, amplifiers, and distribution facilities with, or in the immediate vicinity of the associated transmission equipment. The common portion of the carrier supply for three transmitting and three receiving multiplex bays is located in the first transmitting bay. This is shown in Fig. 15.

Since channel banks are located in the equipment area of the office and the group and supergroup bays of the transistor multiplex in the maintenance or patching area, separate, independent harmonic generators have been provided. This permits the elimination of the major part of the interbay cabling and replaces it with intrabay wiring, resulting in large savings in the cabling costs as well as needed relief for

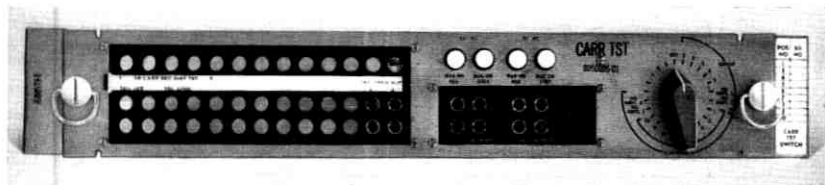


Fig. 17 — Carrier test panel.

offices with cable congestion problems. A cable comparison between the old and the new carrier supply is shown in Fig. 16.

The carrier and pilot supply equipment required for the group and supergroup multiplex bays is divided into the following units: intermediate,* group, and supergroup. Each unit is factory assembled, wired, and tested and mounted as part of the shop-wired bay, thereby reducing costly installation effort. Each design has been patterned specifically for use in floor-saving back-to-back bay lineups. Active equipment is connected into the bay circuits on a plug-in basis which permits rapid service restoration and planned growth with minimum investment in equipment for the initial channels.

In addition to the inherent reliability obtained through careful system design and selection of high quality components and connectors, alternate units are automatically switched into service in the event of a failure affecting more than 360 channels. Consistent with the current protection philosophy, for circuits common to a lesser number of channels, monitored spare units are provided as a standard part of the equipment.

The extensive use of in-service testing as well as convenient test facilities have greatly eased the maintenance burden. An example is a new carrier test panel, shown in Fig. 17, which permits measuring all carriers on a switch basis. Supplementary outlets are also provided to obtain a 64- and 92-ke frequency for test set calibration.

Up to this point the carrier supply equipment described has been primarily for the groups and supergroups. Of equal importance is the channel carrier supply, which has also been miniaturized to the extent that it is now a part of the channel bank bays. A standard three-bay layout is shown in Fig. 18.

This new arrangement includes the channel bank equipment, carrier supply, and fusing for 360 channels and facilities for mounting the primary frequency supply (in place of one channel bank) when required. The channel banks are the transistor A5 type, which have been in regular

* The intermediate supply is here taken to include the pilot supplies.

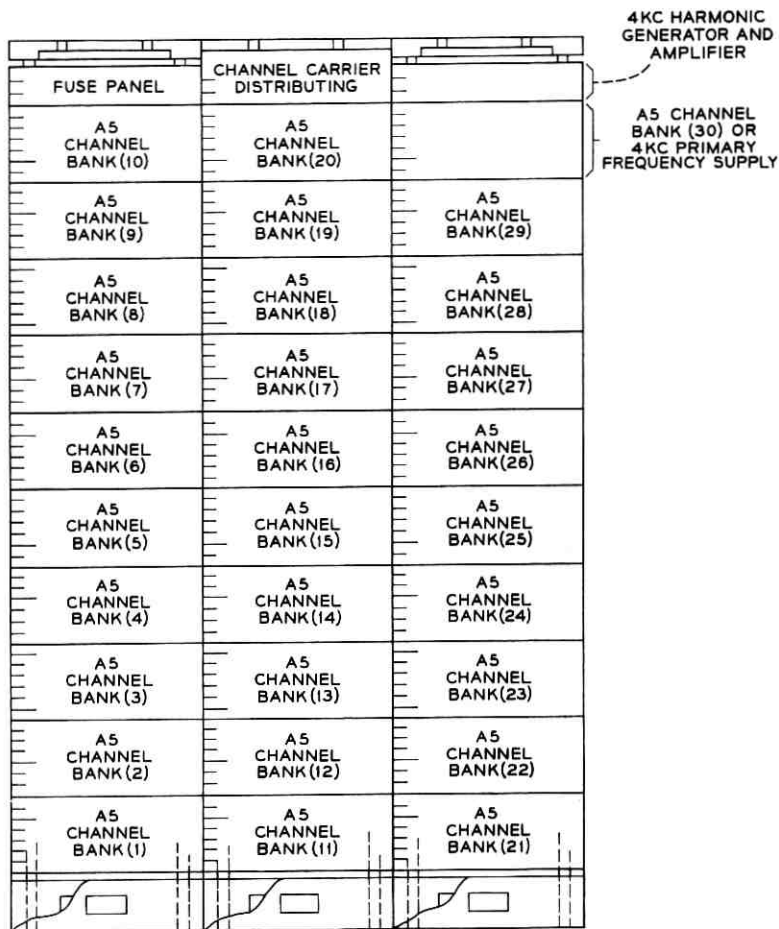


Fig. 18 — Arrangement for 30 A5 channel banks and associated carrier supply.

production since the Fall of 1960.⁹ By dispersing the carrier supply over two bays and increasing the fuse panel capabilities, ten channel banks are mounted in each bay. Since a majority of the connections are voice connections to 4-wire patch bays and since channel banks are often provided singly, channel bank bays are not shop-wired. The channel modem units are removable. Consistent with this and the over-all multiplex design philosophy, active units of the channel carrier supply are plug-in. This supply is independent of the group and supergroup carrier supply, generating and distributing all channel carriers locally.

To summarize, fourteen 11'-6" bays of carrier supply equipment were

formerly required to provide channel, group, and supergroup carriers for 1860 channels. If the new transistor carrier supply were concentrated it would occupy slightly over one-half a bay. The part this plays in the future growth and planning for L multiplex facilities is quite apparent. Often miniaturization has resulted in higher equipment costs with the counter-balance being the floor space saved. Such is not the case with the multiplex equipment where lower unit costs were also achieved.

6.3 *Apparatus Mounted in Plastic*

The various equipments that comprise the carrier supply have been packaged in a manner consistent with the approach used in the associated transmission equipment. However, because of the wide range of carrier and pilot frequencies and the power requirements, some design latitude was necessary. A further design consideration was the relatively low demand of most of the carrier supply units as compared to most of the transmission units. This generally eliminates the use of mass production techniques since they become inefficient and therefore uneconomical in such situations. To counteract this, a process developed by the Western Electric Company and nicknamed "AMPLAS" was selected for much of the packaging.

Fundamentally, this process consists of hand- or machine-inserting components, by their pigtailed, into a soft CAB* mold. A thermosetting epoxy material is poured into the mold and hardened. The CAB mold is stripped away from the assembly, leaving the epoxy board holding the components. Pencil wiring is used for interconnections with each connection individually soldered; however, mass soldering techniques are contemplated for the future.

Figs. 19 and 20 show the apparatus and wiring sides respectively of a typical carrier supply board. Based on experience to date, this packaging method has proven to be very economical.

In view of the similarity in construction of many of the specific carrier supply designs it would serve no useful purpose to describe them all; therefore, only a representative sample will be discussed in more detail.

6.4 *Group Carrier Supply*

The function of this unit is to generate, amplify, and distribute the five group carriers to secondary or local bay distributing facilities which in turn supply group carriers for 1860 channels. A photograph is shown in Fig. 21.

This is a completely shop-assembled, wired, and tested package. All

* Cellulose Acetate Butyrate.

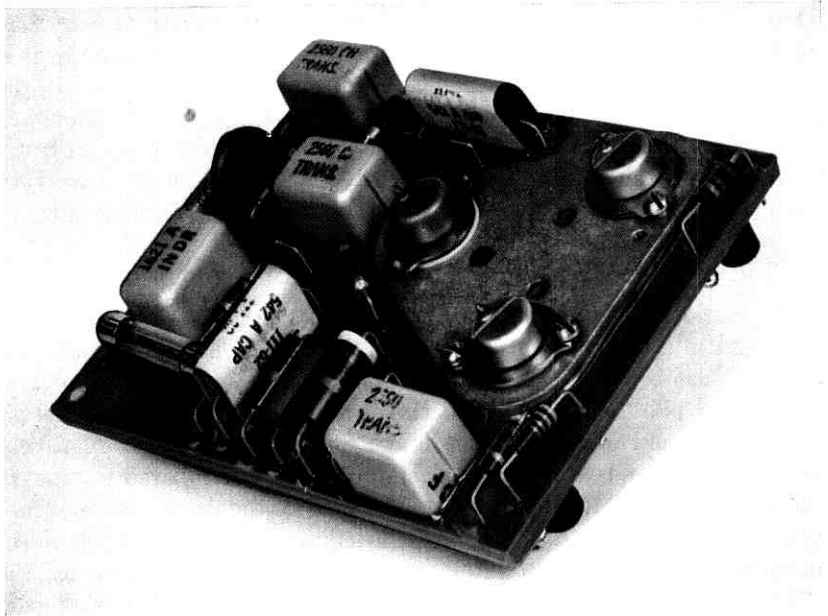


Fig. 19 — Amplas board — apparatus side.

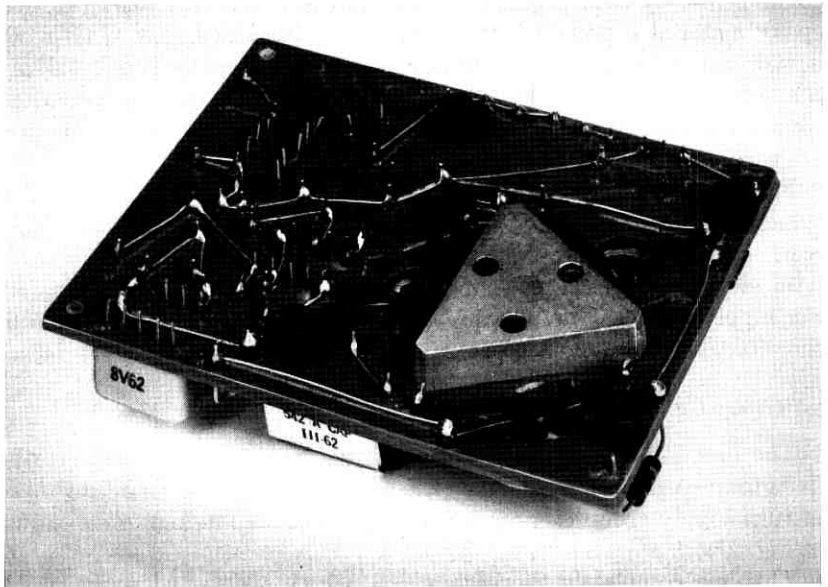


Fig. 20 — Amplas board — wiring side.

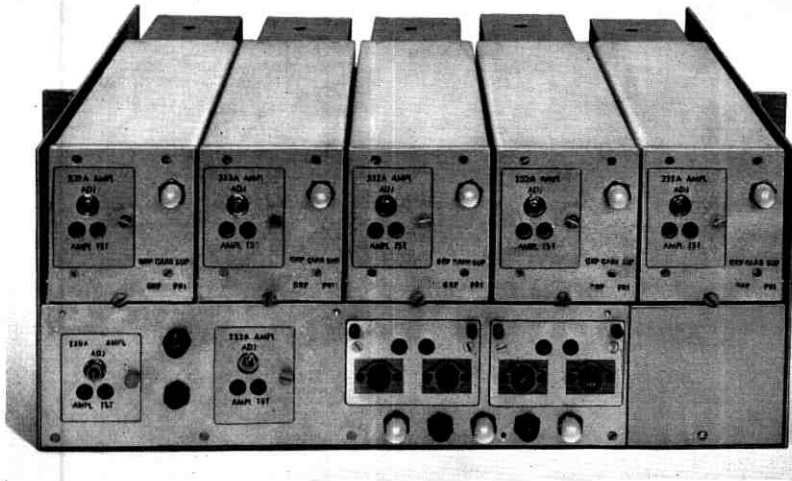


Fig. 21 — Group carrier supply.

active units are plug-in and in the case of the 12-kc amplifiers alternate units are automatically switched into service in the event of a failure. The lower shelf of this two-tier assembly contains a drawer-type chassis that includes the two 12-kc amplifiers, harmonic generator, alarm boards, and monitored spare distribution amplifiers. This arrangement provides for relatively easy replacement of even the passive components. Visual alarm indications previously described are provided on the front face to aid in making a quick determination of equipment troubles.

The upper shelf contains the five group distribution modules. Each module, complete with filter, amplifier, and distribution bus, is connected into the circuit by means of a plug. In this instance, the plug-in feature was used to facilitate the intra-panel wiring with the ability to replace defective modules coming as an added benefit.

A view of a module with the cover removed is shown in Fig. 22. Note the open chassis, which permits all shop connections to be made with a minimum of effort. The previously described AMPLAS-type board is used for the distribution bus.

Each module contains a 232A amplifier, shown in Fig. 23. This amplifier is used in any of the five group modules and can be replaced without removing the complete module. The rather large heat sink provided for the power transistor is connected directly to the front face in order to transfer heat quickly to the front radiating surface.

The complete group carrier supply is provided as a unit for the L600A

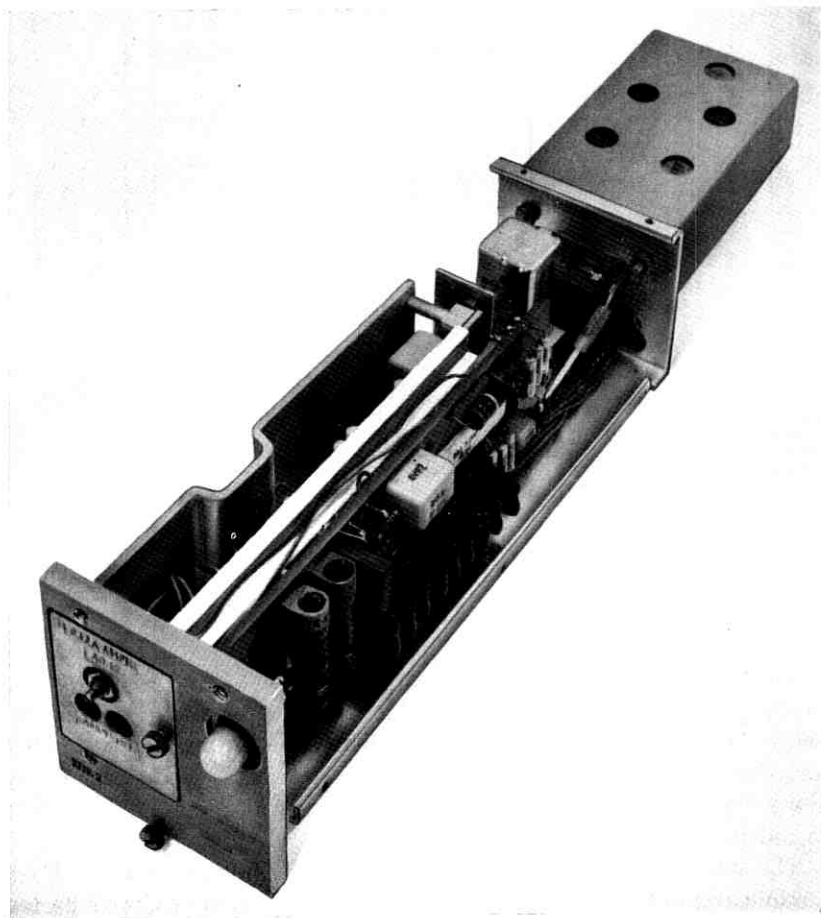


Fig. 22 — Group carrier primary distribution module.

multiplex bay; however, it is possible to provide a partial supply if desired by merely omitting modules.

6.5 *Supergroup Carrier Supply*

The design of the supergroup carrier supply presented one of the more challenging equipment undertakings since, in addition to miniaturization, it was necessary to have maximum flexibility for the ten-supergroup L600A application as well as the thirty-one supergroup L1860A application.

The complete supergroup carrier supply unit generates, amplifies, and distributes the supergroup carriers directly to the modulators and

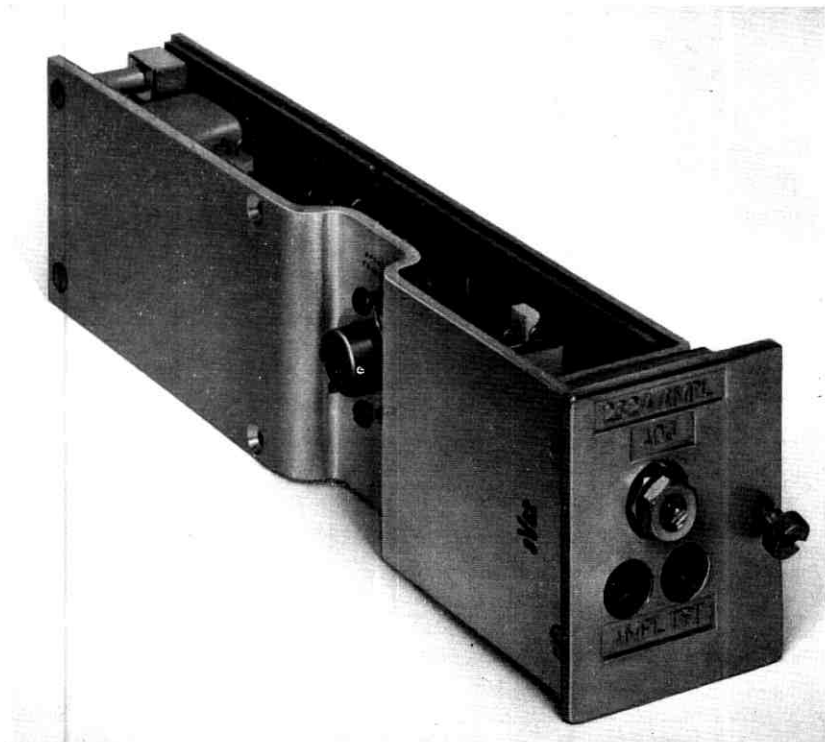


Fig. 23 — 232A amplifier.

demodulators. The shelf arrangement, shown in Fig. 24, is capable of providing all the necessary supergroup frequencies for any standard L600 application. The frequencies not required for a particular circuit arrangement need not be provided, but they can easily be added at a future date with a minimum of installation effort.

The three shelves are designated A, B, and C. The A shelf contains the 124-kc harmonic generator and the distribution modules for two carrier frequencies. The B and C shelves are primarily for mounting the remaining regular and spare distributing modules. For L1860A, a different A shelf containing the 80- and 124-kc harmonic generators and one distributing module is provided. Also a fourth shelf is added to provide distribution facilities for the four additional carriers required to directly form the upper supergroups.

Each shelf is a separate shop-wired unit with outside connections made through miniature coaxial plugs and jacks. Following the general multiplex pattern the amplifiers are plug-in, with harmonic generator

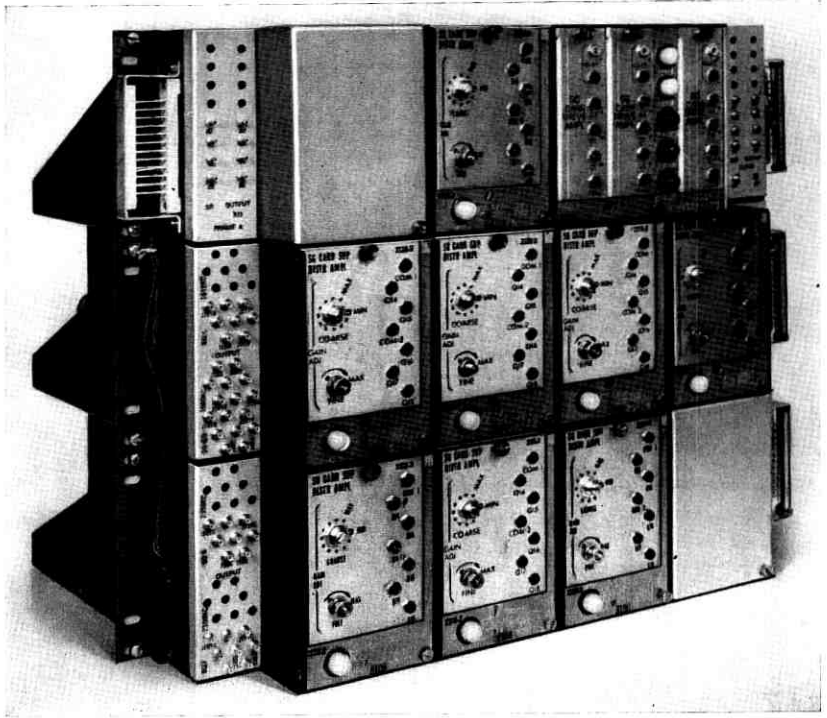


Fig. 24 — Supergroup carrier supply.

amplifiers protected on an automatic switch basis. The distribution amplifier, shown in Fig. 25, is identical for all frequencies; in fact, it contains the preamplifier and power amplifier on a common frame. The three transistors of the power amplifiers are mounted on a single triangular aluminum block which is in solid contact with the chassis to facilitate heat transfer.

All test outlets, adjustments, and output connections are located on the front surface for ease of maintenance. Each module contains a visual alarm indication to aid in quickly locating a defective unit. As the installation grows, additional outputs serving up to 1860 channels may be obtained from connectors located on the side distribution panels.

6.6 Channel Carrier Supply

The channel carrier supply is comprised of two separate major units: the channel harmonic generator and filter unit and the channel carrier distribution unit. In an ideal arrangement these will be located at the

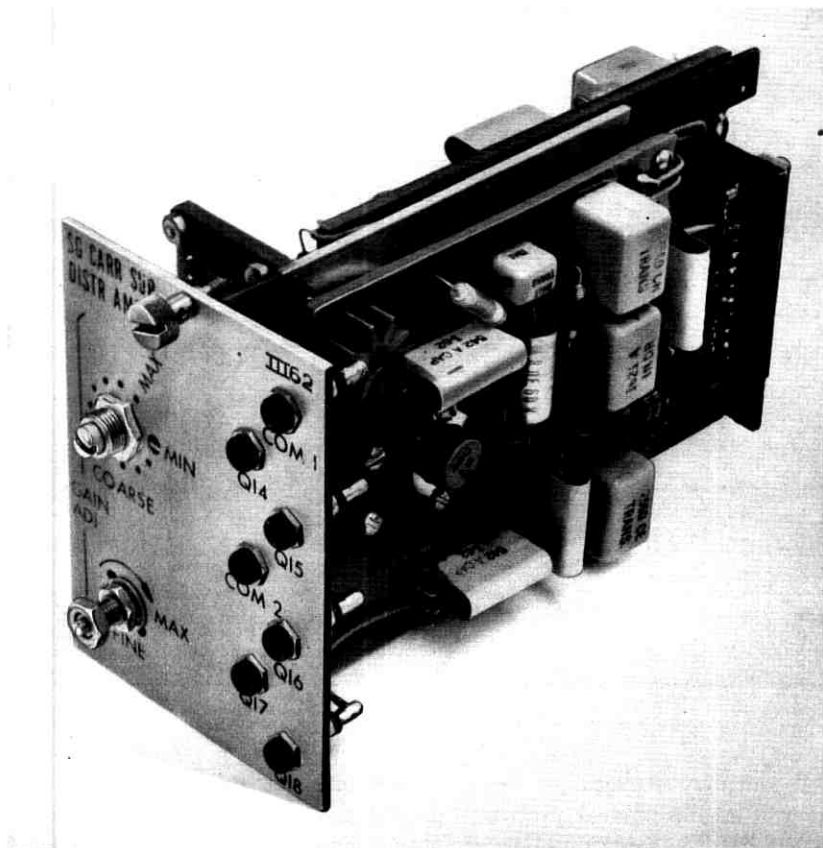


Fig. 25 — Supergroup carrier distribution amplifier.

top of adjacent bays. The channel carrier supply, as mentioned earlier, is located with the channel banks in the equipment area.

The generator and filter assembly departs from the module arrangement and follows the door-type structure of the A5 channel banks. It contains the amplifier, harmonic generators and filters necessary to supply channel carrier frequencies to 360 channels. The working amplifier is plug-in, and space is available for a spare monitored amplifier. Currently it is recommended that one spare amplifier be provided for each five regulars. All maintenance is from the front since these bays may be located in back-to-back lineups.

The distribution unit presented a design problem which is becoming increasingly common in this age of miniaturization when the component size reductions are not followed by equivalent cable reductions. This

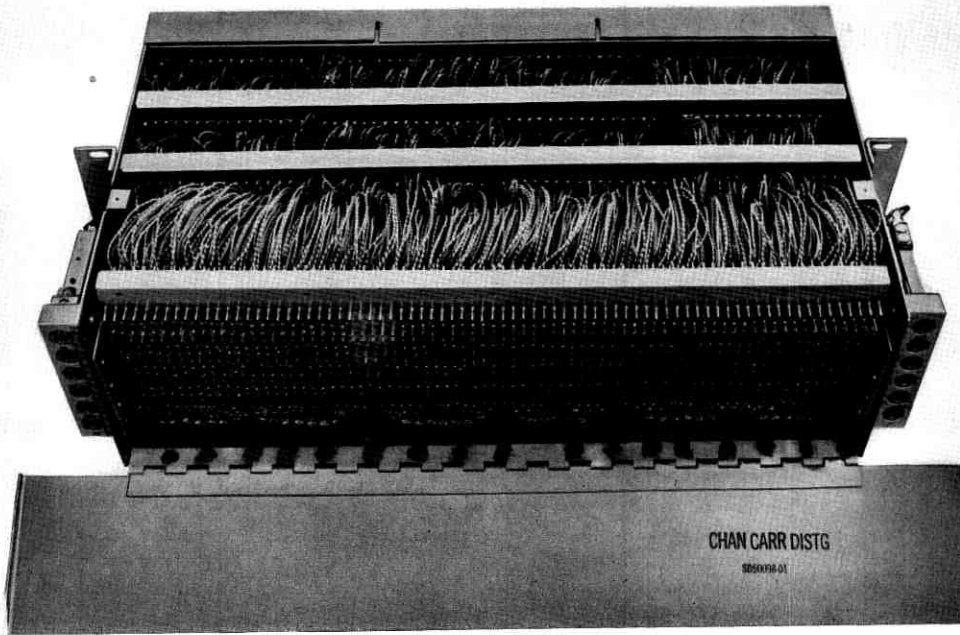


Fig. 26 — Channel carrier distribution unit.

will continue to plague us, because cable size reductions involve circuit as well as mechanical problems. In the distribution assembly, shown in Fig. 26, the same number of channel modems as in previous designs is supplied; however, the old design required about one bay of mounting space as opposed to 12 inches for the new design.

Means have been provided for the installer to connect 360 shielded pairs of wire to a terminal strip which terminates all distribution bus outputs. A screw-type adjustment similar to that used on the A5 channel bank and E6 repeater is available to facilitate adjusting the bus level as additional taps are used. Formerly this required a soldering operation. An additional feature is the jacks provided on the front of the panel for measuring the levels of the twelve distribution busses.

VII. CONCLUSION

A laboratory model of the carrier supply has been on field trial at Dallas, Texas, since March, 1962, with satisfactory results. The first commercial L600A installation, on a TJ microwave route in New Jersey, was cut into service in sections from July through September, 1962.

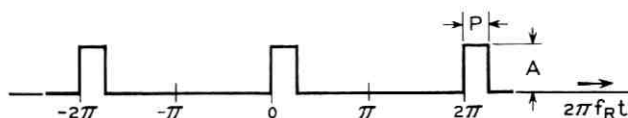


Fig. 27 — Pulse train.

VIII. ACKNOWLEDGMENTS

The system described here, as with the rest of the multiplex, is the result of the efforts of many people in several departments of the Bell Telephone Laboratories; specifically Systems Engineering, Systems Development and Device Development. Although unable to cite all those who participated in this design, the authors wish to acknowledge particularly the substantial contributions of F. C. Kelcourse and T. J. Haley in the circuit design area, and M. F. Stevens and L. F. Travis in the equipment design area, in addition to the authors of the companion papers.

APPENDIX

Selection of an appropriate pulse width for the harmonic generator circuits was investigated early in the design of the new multiplex carrier supply. In the old carrier supply this choice was somewhat restricted since the 4-ke harmonic generator supplied all of the channel and group carriers and was chosen primarily on the basis of generating sufficient power for the channel carrier supply, thus eliminating the need of channel carrier amplifiers. As a result, group carriers exhibited a negative slope-versus-frequency characteristic.

With the decentralization of the carrier units in the new multiplex and the specialized harmonic generators for the channels, groups, and supergroups, different pulse widths could be selected to best fit each application.

If one considers a continuous series of pulses as shown in Fig. 27* with a period $T = 1/f_h$, amplitude A and pulse width P and, further, that each pulse is perfectly rectangular, one can completely specify the harmonic content of any given pulse width P in the frequency interval of interest. From Fourier analysis, one can then define the wave function as:

* The pulses feeding the "even" channel carrier filters may be represented directly by the described pulse train. Other pulses in the carrier supply may be synthesized by linear superposition of two pulse trains with amplitudes of opposite signs.

$$f(t) = \sum_{n=1}^{\infty} (a_n \cos n\omega t + b_n \sin n\omega t)$$

where

$$a_n = \frac{A}{n\pi} \sin nP \quad (P \text{ in radians at } f_R)$$

and

$$b_n = \frac{A}{n\pi} (1 - \cos nP).$$

TABLE I

(a) Channel and Intermediate Supply
Amplitude of Harmonic Relative to Fundamental Pulse of
Unit Amplitude and Frequency of 4 kc

Pulse Width (μ s)		1	2	3	4
Harmonic Frequency (kc)	64	0.0079	0.0155	0.0225	0.0286
	108	0.0078	0.0148	0.0200	0.0231

(b) Group Carrier Supply
Amplitude of Harmonic Relative to Fundamental Pulse of
Unit Amplitude and Frequency of 12 kc

Pulse Width (μ s)		1	2	3	4
Harmonic frequency (kc)	420	0.0176	0.0088	0.0132	0.0154
	612	0.0117	0.0080	0.0062	0.0123

(c) Supergroup Carrier Supply
Amplitude of Harmonic Relative to Fundamental Pulse
of Unit Amplitude and Frequency of 124 kc

Pulse Width (μ s)		0.1	0.5	1.0	2.0
Harmonic Frequency (kc)	1116	0.0241	0.0694	0.0241	0.0454
	1612	0.0237	0.0281	0.0460	0.0316
	2108	0.0230	0.0056	0.0110	0.0210
	3100	0.0210	0.0027	0.0053	0.0104

In terms of the pulse parameters, the peak amplitude of any harmonic "n" is given as:

$$A_n = \frac{2A}{n\pi} \sin n\pi f_R P \quad (P \text{ in units of time consistent with } f_R).$$

Table I shows an evaluation of the above expression for various pulse widths for both the channel (and intermediate), the group, and the supergroup carrier supplies. The numbers in Table I represent the relative amplitudes of the harmonics with respect to the pulse amplitude "A."

For the channel carrier supply a pulse width of 4 μ sec was chosen as a good compromise between slope and amplitude. For the group supply 1 μ sec was selected and 0.1 μ sec was used for the supergroup carrier supply.

REFERENCES

1. Ehrbar, R. D., Elmendorf, C. H., Klie, R. H., Grossman, A. J., et al., The L3 Coaxial System, B.S.T.J., **32**, July, 1953.
2. Roetken, A. A., Smith, K. D., and Friis, R. W., B.S.T.J., **30**, October, 1951, p. 1041.
3. Caruthers, R. S., The Type N1 Carrier Telephone System: Objectives and Transmission Features, B.S.T.J., **30**, 1951, p. 1.
4. Gammie, J., and Hathaway, S. D., The TJ Radio Relay System, B.S.T.J., **39**, July, 1960, p. 821.
5. Kinzer, J. P., Laidig, J. F., et al., The TH Microwave Radio Relay System B.S.T.J., **40**, 1961.
6. Hallenbeck, F. J., and Mahoney, J. J., Jr., B.S.T.J., this issue, p. 207.
7. Graham, R. S., Adams, W. E., Powers, R. E., and Bies, F. R., B.S.T.J., this issue, p. 223.
8. Clark, O. P., Drazy, E. J., and Weller, D. C., B.S.T.J., this issue, p. 319.
9. Blecher, F. H., and Hallenbeck, F. J., The Transistorized A5 Channel Bank for Broadband Systems, B.S.T.J., **41**, Jan., 1962, p. 321.

A Phase-Locked Primary Frequency Supply for the L Multiplex

(Manuscript received October 30, 1962)

By O. P. CLARK, E. J. DRAZY, and D. C. WELLER

The carriers and pilot tones of the L multiplex are all derived from a single source of 4 kc. The new primary frequency supply provides this primary frequency. It is designed to operate phase-locked to a pilot tone derived from an incoming carrier system and hence to have no frequency error with respect to that tone. The unit features a high degree of frequency stability, even in the absence of the pilot; it features more reliable performance and lower maintenance as well as marked size reduction compared with the older equipment.

I. INTRODUCTION

In single-sideband suppressed-carrier transmission systems, of which the L-type multiplex is a modern example, it is necessary to supply, at each terminal, sequences of harmonically related carrier frequencies to power the various modulators, and also to provide line pilot frequencies for system regulation. Because the accuracy requirements for these frequencies are severe, it has proven most economical and technically satisfactory to derive them all from a single highly accurate and stable source which, in the case of the L-type multiplex, has been named the primary frequency supply. The means by which the ensemble of carrier and pilot frequencies is derived therefrom are described in a companion article.¹

II. REQUIREMENTS — GENERAL

To provide flexibility within and promote orderly growth of the communications plant, it is desirable that newly designed equipment operate compatibly with and serve as electrical replacements for older counterparts, where such exist. This compatibility must, however, be achieved without limiting the performance capabilities of the new equipment to those representative of the older. The predecessor of the primary fre-

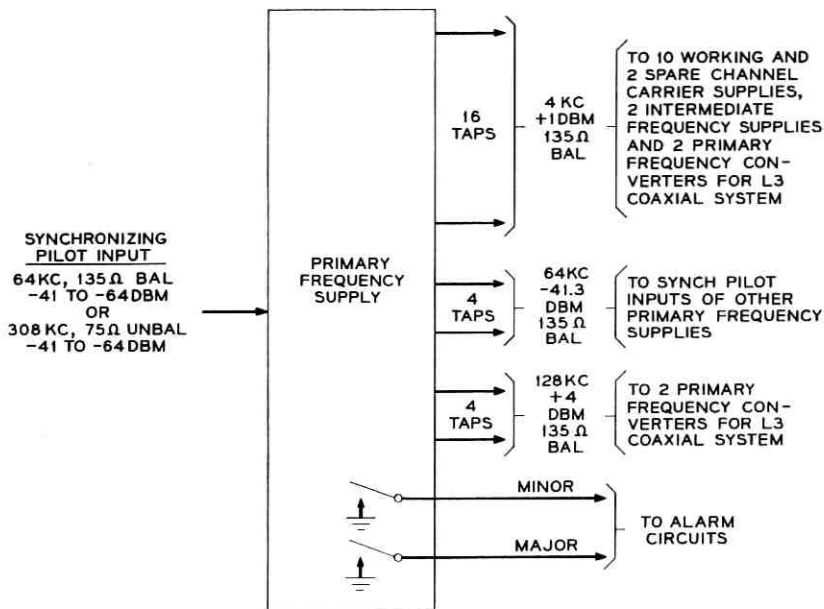


Fig. 1 — Primary frequency supply interfaces.

frequency supply was the 4-kc supply of the L telephone terminal,² together with various physically separate, but functionally associated synchronizing and distributing circuits. Since it was evident that there were many installations where combinations of old and new equipment would occur, compatibility assumed great importance. Fig. 1 shows the electrical interfaces between a fully loaded primary frequency supply (as well as its functional predecessor) and the remainder of the multiplex equipment. Three output frequencies are provided, (i) the basic 4 kc, corresponding to the channel separation of the multiplex,* (ii) 128 kc, required for carrier generation in associated L3 coaxial system terminals, and (iii) 64 kc, which may optionally be used for synchronization, as described later. All carrier and pilot frequencies of both the L-type multiplex and the L3 coaxial system are integral multiples of the basic 4 kc.

With compatibility and replaceability postulated, *minimum* performance requirements, equivalent to those of the older equipment, are established. At this point, the execution of a prudent design requires consideration of probable future applications of the equipment which might increase the severity of these requirements, and, where economically

* Alternatively, some of the 4-kc output taps may be used to supply a reference frequency for carrier generation in other carrier systems.

feasible, provision either of performance consistent with these anticipated applications, or means for future improvement without complete redesign. Such anticipated applications include, among others, increasingly sophisticated systems for data transmission, stereophonic program transmission, and establishment of a system-wide frequency standard. In addition, the rapid growth of intercontinental toll facilities via submarine cable and, as seems likely, via satellite-borne systems, demands conformance with accepted international standards of performance.³

In addition to satisfying the performance requirements arising from the considerations outlined above, the design of the primary frequency supply must comply with constraints of space, environmental conditions, and available power supply, and, because of the large number of dependent message channels — as many as 3600 — must guarantee a very high order of reliability and provide easy access for maintenance and repair.

2.1 *Frequency Accuracy*

The frequency accuracy requirements of the primary frequency supply are of two kinds, absolute and relative. Absolute accuracy is a statement of the difference between the nominal and actual frequencies of a given supply measured in terms of real time; relative accuracy is a statement of the difference in actual frequencies of two supplies. Both requirements must be stated in terms of allowable drift per interval of time; for the moment this will be loosely defined as the normal maintenance interval, the duration of which will be determined later. Since all derived frequencies are generated by frequency multiplication, the fractional accuracy requirement for any carrier or pilot applies also to the primary source.

The absolute requirement stems from the necessity for a proper "fit" of the derived carrier and pilot frequencies to the transmission characteristics of the highly selective filters to which they are applied. Of these frequencies, the most critical is the 92-kc group regulating pilot,⁴ not only because of the high selectivity of the corresponding filters, but also because of the serious effects of such errors on system performance. The tolerable error at this frequency has been established at one cycle per second. Rounded off, this is one part per 10^5 per maintenance interval.

Relative frequency accuracy is the measure of the difference between the primary frequency and its derivatives within a given multiplex terminal to those within its associated counterpart at the far end of a

transmission facility. The consequence of inaccuracy of relative frequency is frequency translation error, in which all components of a message are displaced in frequency by an equal number of cycles, with resultant destruction of their normal harmonic relationship and hence, loss of fidelity of the recovered information.

The channel most susceptible to this type of degradation is that which occupies the highest line frequency produced by the L3 coaxial terminal when it is used in conjunction with (and with its carriers derived from) the L multiplex. The corresponding "virtual" carrier frequency is 8.284×10^6 cycles per second. The more susceptible of present services, such as program and voice-frequency telegraph, suffer impairment due to frequency errors in excess of two cycles per second. These considerations lead to a relative frequency accuracy requirement of 2 parts per 8.284×10^6 , or, rounded off, 2 parts per 10^7 per maintenance interval. Comparison of this requirement with that for absolute frequency accuracy reveals that the relative accuracy requirement is the more severe by more than an order of magnitude, and is therefore controlling.

In order to satisfy the relative accuracy requirement without correspondingly increasing the absolute requirement to values which were economically impractical at the time of the design of the older 4-kc supply, so-called "synchronous" operation was adopted. This was accomplished by designating a selected supply as the "master"* and transmitting, as a pilot, an integral harmonic (64 kc or 308 kc) of its output frequency. At controlled terminals the corresponding locally generated harmonic is compared with the incoming pilot, and correction automatically effected. In the older L-type telephone terminals, the correction is via an electromechanical servo loop. Although this system is theoretically capable of perfect correction, its sluggish nature, together with the effects of static friction, result in a usual relative error of about two parts per 10^7 , which is marginal.

It is important to remark here that multiplex terminals often receive synchronizing and message information over different transmission paths. Increasing emphasis on system security points up the desirability of minimizing service degradation in those cases where the synchronizing path may be interrupted while the message path remains intact. In the older 4-kc supply, the static friction of the mechanical elements

* At present, the "master" supply is located in New York City; actually, it is in turn controlled by the Bell System Primary Standard of Frequency at Murray Hill, N. J., the accuracy of which is one part in 10^9 , indefinitely (achieved by periodic correction).

From this master are controlled very nearly all of the multiplex terminals in the continental United States. As many as nine synchronizing links may be interposed between the master and the most remote terminal.

of the synchronizing servo has provided a useful "memory" to prevent any substantial change of frequency during such interruptions.

The new primary frequency supply is, therefore, required to have the capabilities both of synchronizing and of being synchronized by existing 4-ke supplies, and of satisfactory operation on a free-running basis for appreciable intervals of time. Despite the value of the frictional "memory" of the electromechanical servo of the 4-ke supply, maintenance problems led to its abandonment in favor of an all-electrical servo which, during normal operation, maintains a constant phase relationship between the locally generated and incoming pilots, and so completely eliminates relative frequency errors. The free-running accuracy requirement is satisfied by making the free-running frequency accuracy of each primary frequency supply sufficiently great (and, therefore, the required frequency-control range so small) that the requirements of the preceding section are fulfilled even though control is relaxed due to pilot interruption. This requirement becomes 1 part in 10^7 when allowance is made for two oppositely drifting supplies. Present state-of-the-art permits quite economical realization of this accuracy for one-month intervals; at somewhat greater cost the interval can be extended to six months. The requirement, therefore, was established as lying between the above limits, taking the following into account:

1. Relatively few supplies will be loaded with full 1800-channel systems; in these cases it appears most economical to shorten the maintenance interval, rather than burden all systems with the added cost of unneeded accuracy.
2. The probability of the most susceptible services' occupying the least favorable carrier frequency slot is small.
3. The possibility of "network" interconnection of synchronizing pilot is being studied. If adopted, this will make the probability of pilot failure almost negligible.

To meet this objective, two distinct phase-locked frequency supply circuits are provided, either of which supplies the useful output at any given time. A minimum of components — all passive — appear in the common path following the combining circuits. The latter employ hybrid circuitry, and are so designed that either input can be open- or short-circuited without materially affecting transmission from the alternative supply via the conjugate input. Failure of the working supply initiates a transfer to the alternate supply, and simultaneously provides a minor alarm and visual indication so that repairs can be effected. The phase-lock of the two supply circuits, together with proper phasing of the hybrid transformers in the combining circuits and the high switching speed all help to minimize transmission disturbances due to the

transfer. Minor alarms and visual indications are also provided to indicate failure of the non-working supply circuit, failure of the incoming pilot, and end-of-range of the automatic frequency and phase-control circuits. The latter function is performed by meter-relays, which also serve as indicators during adjustment of the centering of the phase control range. Complete failure of any of the three output frequencies results in a major alarm.

Routine maintenance consists of observing and adjusting to a minimum the voltage applied to the varactor diode to effect phase-lock. This has the effect of matching the natural frequency of the controlled oscillator to the frequency of the controlling master oscillator and hence minimizing the frequency deviation that would result were pilot control lost. One-month maintenance intervals are expected to be sufficient for newly-installed equipment, with increases to three or six months after aging of the oscillator crystals.

Test jacks are provided to permit the oscilloscopic examination of significant waveforms in the digital circuitry. To correct malfunctioning circuitry, the plug-in circuit containing the defective elements will be removed and replaced by a spare. Field repair of defective plug-in circuits will not be attempted.

2.2 *Mechanical, Environmental and Power Supply Requirements*

The primary frequency supply is designed to mount in a standard nineteen-inch relay rack, in a space $5\frac{1}{4}$ inches high by 15 inches deep. The ambient temperature limits are 38° to 120° F. Performance must be unaffected by a mechanical shock having 3-g acceleration. Available power supply is -21 to -28 vdc.

III. TIMING OSCILLATOR

The basic timing oscillator of the primary frequency supply is crystal-controlled and operates at a frequency of 1.024 mc. This choice of frequency is, as required, a binary multiple of 4 kc, and is high enough to permit use of a crystal cut which is favorable to stability, yet not so high as to require an inordinate number of frequency dividers. A modified Pierce circuit⁵ is used, as shown in Fig. 2. The frequency of the oscillator is controlled, over a range of ± 25 parts per 10^7 , by the capacitance of the varactor diode, which is effectively in series with the quartz plate. The capacitance of the varactor diode is, in turn, governed jointly

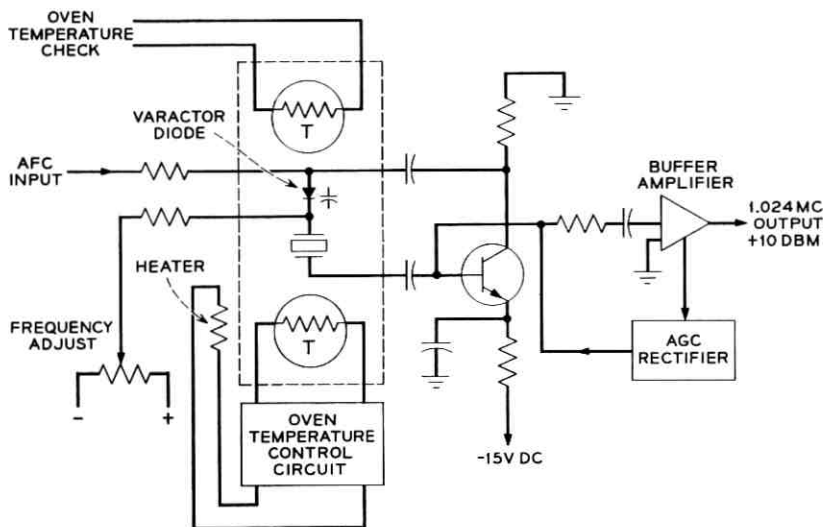


Fig. 2 — Timing oscillator.

by the frequency-control voltage and the setting of the frequency adjustment potentiometer. The five-stage buffer amplifier isolates the oscillator from the connected load, and permits the oscillator proper to operate at a low level, which is favorable to stability.

The crystal is a plano-convex AT cut quartz plate, operating in the fundamental mode. It is mounted in an evacuated cold-welded metal enclosure and, together with the varactor diode, is housed in an oven which is provided with a stepless oscillating closed-loop temperature control system which maintains the temperature constant within $\pm 0.15^\circ\text{C}$ for all anticipated ambient conditions. A monitoring thermistor is provided so that the oven temperature can be checked at any time.

IV. DIGITAL FREQUENCY DIVIDERS

Although the inputs and outputs of the primary frequency supply are signals of sinusoidal waveform, the advantages of digital techniques for the necessary frequency division, switching, alarm, and frequency and phase comparison functions led to their adoption for the internal circuitry. Among these advantages are:

1. Economies resulting from the use of identical digital circuits for various logic functions.

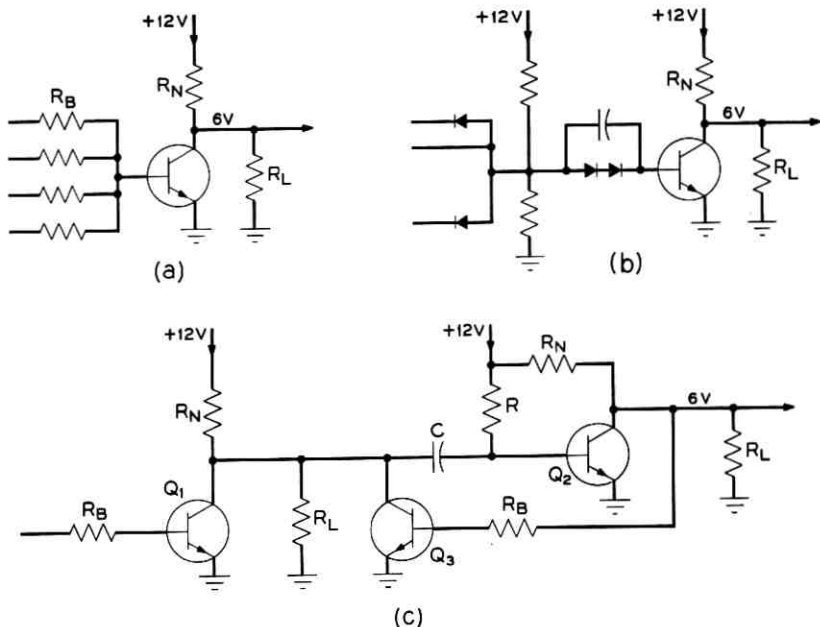


Fig. 3 — Logic circuits.

2. The relative ease with which the 77-to-1 division, necessary for operation with a 308-kc pilot, could be achieved.

3. The go-no-go nature of digital circuits, which contributes to ease of maintenance.

4. The inherently excellent time precision which can be realized with this type of circuitry.

The three basic transistor logic circuits used in the frequency divider units are shown in Fig. 3. A 12-volt supply with a transistor node resistance (R_N) of 2370 ohms is used for all logic functions. The value of the load resistor R_L is selected to limit the maximum collector-to-emitter voltage to 6 volts. The majority of logic functions are performed by the Transistor-Resistor-Logic (TRL) of Fig. 3(a). When an increase in switching speed is required from this circuit, a capacitor is used in parallel with the base resistor R_B . Since silicon transistors are used, reverse bias for the base circuit is not required. The Transistor-Diode-Logic (TDL), Fig. 3(b), is used to perform fast gating functions where as many as eight input signals are received. The two series diodes in the base circuit develop the required reverse bias to insure turnoff of the transistor.

This method of obtaining bias proves more efficient than providing a special power supply when relatively few TDL circuits are required as in this equipment. The Univibrator (U) logic circuit, Fig. 3(c), is used to generate a pulse of controlled length from a positive-going pulse of any duration greater than 0.1 microsecond. The length of the pulse is determined by the time constant of R and C.

The signals received by the digital frequency dividers from both the timing oscillator and the pilot amplifier are sine waves. It is necessary to convert these signals to logic type signals before they can be used. This conversion must be accurate within 0.1 microseconds. The sine wave to logic pulse converter used to perform this function is shown in Fig. 4. Q_1 is an emitter follower, with temperature stabilization of its bias circuit provided by CR_1 and CR_2 . Q_2 provides high switching current to Q_3 , which is operated as a common-base amplifier. The base voltage E on Q_3 is set to a value that yields a symmetrical square wave at the collector of Q_3 for a sine wave input at Q_1 . The square wave is converted to logic level by coupling through C to Q_4 with a dc restorer diode CR_3 .

The binary frequency division is performed by a three-stage transistor circuit (see Fig. 5). Q_1 and Q_2 form a flip-flop circuit with diode-capacitor steering for the input signal. The negative edge of an input square wave is used to change the state of the flip-flop circuit. A buffer amplifier, Q_3 , is used to minimize the loading effects of associated logic circuits on the flip-flop counter stage. The maximum delay through the binary counter stage is 0.05 microsecond.

The method of using the divide-by-256 unit with the timing oscillator, pilot amplifier, and digital discriminator to obtain 64-ke pilot phase-

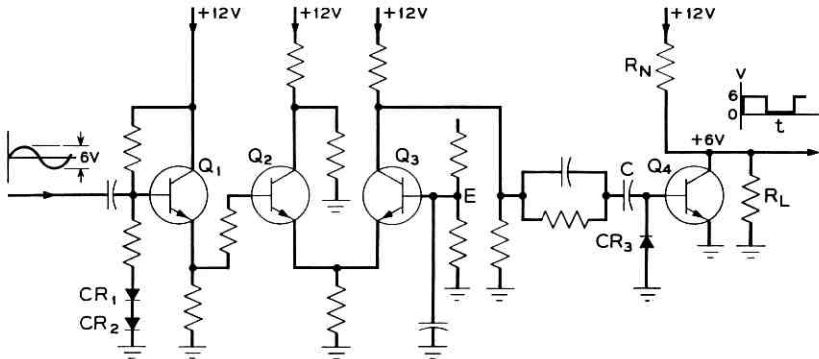


Fig. 4 — Sine-wave to logic pulse converter.

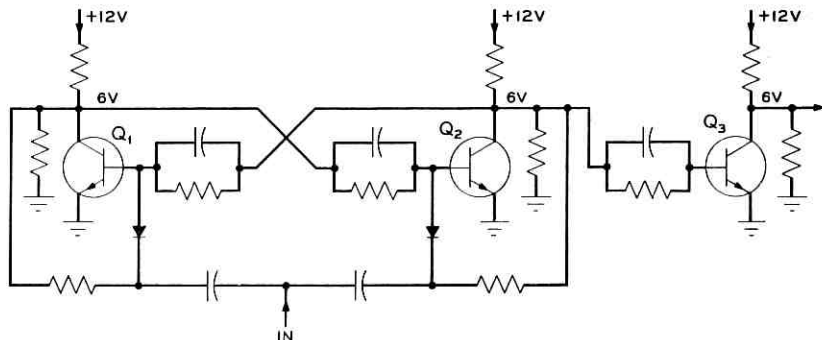


Fig. 5 — Binary counter circuit.

lock, and generate the basic frequencies for the carrier supply, is shown in Fig. 6. The 64-kc "clocked" pulses for the digital discriminator are generated by diode gate no. 1, which receives gating pulses from the first four binary counter stages, and the clocking pulses from the timing oscillator. The digital discriminator⁶ compares the phase of the 64-kc clocked pulses with that of the 64-kc pilot and generates a dc voltage proportional to their phase displacement. This dc voltage is applied

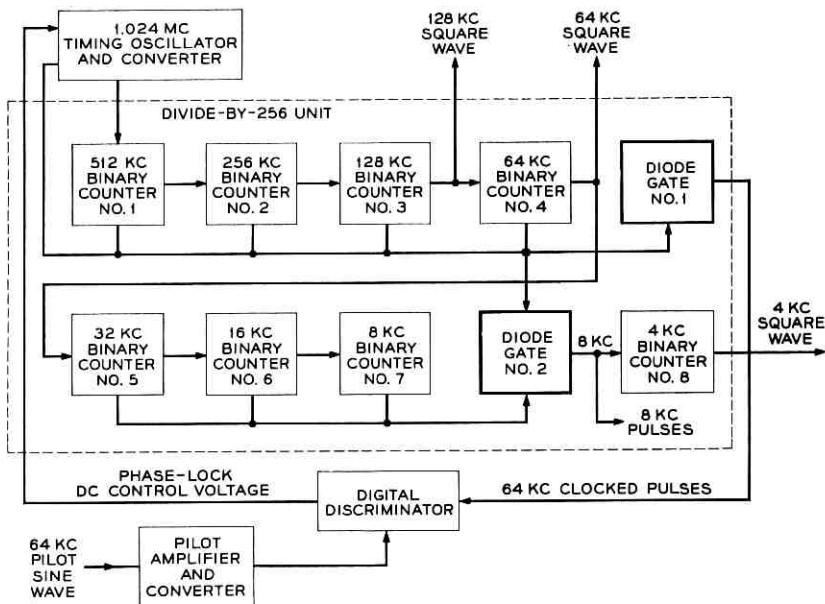


Fig. 6 — Divide-by-256 unit with 64-kc pilot phase-lock connections.

across the varactor diode in the timing oscillator to change the timing frequency for proper phase-lock. Diode gate no. 2 receives gating pulses from the first seven binary counter stages and the clocking pulses from the timing oscillator. The 8-kc clocked pulses generated by diode gate no. 2 are used to drive the eighth binary counter stage and the digital discriminator, when 308-kc pilot phase-lock is required. Driving the eighth binary counter with clocked 8-kc pulses to obtain a 4-kc square wave minimizes the delay variations that may occur through seven binary counter stages.

In those cases where the timing oscillator is to be phase-locked to a 308-kc pilot, it is necessary to use a divide-by-77 unit to generate an 8-kc frequency that can be compared in phase with the 8-kc frequency obtained from the divide-by-256 unit. The method of using the divide-by-77 unit with the timing oscillator, divide-by-256 unit, digital discriminator and 308-kc pilot signal is shown in Fig. 7. The diode gate in the divide-by-77 unit receives gating pulses from seven binary counter stages and a 616-kc clocking pulse generator.

The code for the diode gate to obtain a divide-by-77 is as follows: $64 + 0 + 0 + 8 + 4 + 0 + 1 = 77$. When the 77th pulse enters binary counter number 1, the diode gate detects an "AND" condition and sends a positive pulse to the univibrator. The univibrator immediately (within 0.2 microseconds) initiates the resetting of the seven binary

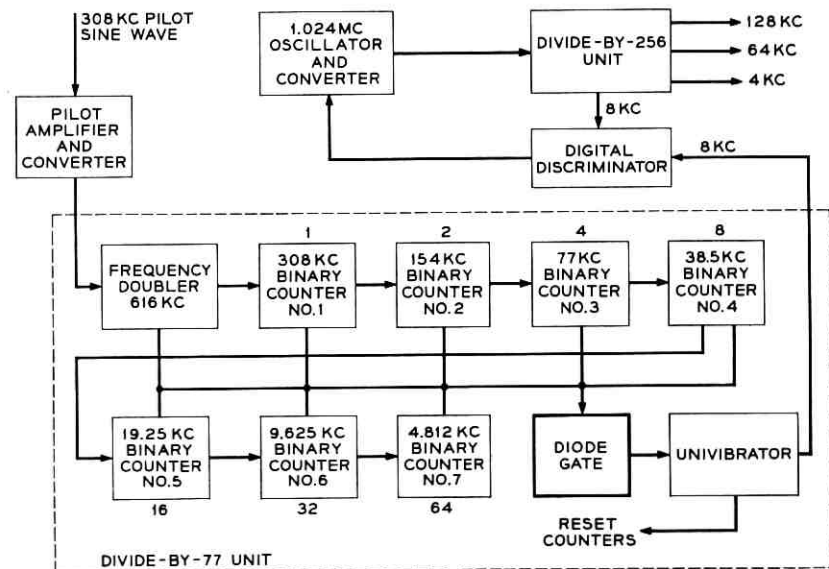


Fig. 7 — Divide-by-77 unit for 308-kc phase-lock.

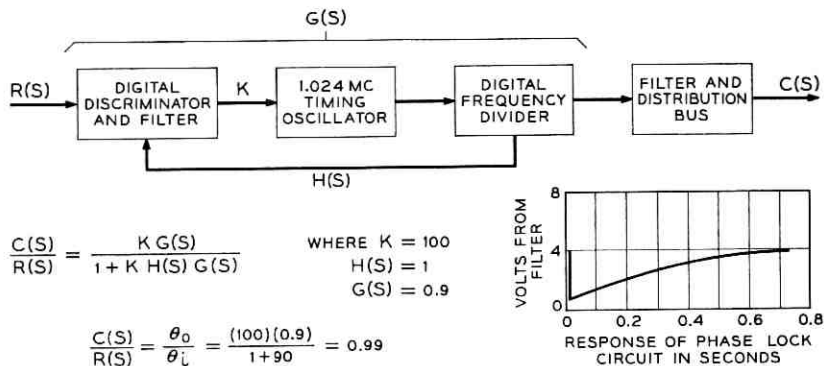


Fig. 8 — Phase-lock control loop.

counter stages to zero. This resetting operation is performed within one microsecond. The 8-ke output from the divide-by-77 unit is used by the digital discriminator to phase-lock the timing oscillator in the same manner described for the 64-ke pilot phase-lock system. Doubling the 308-ke frequency before dividing provides an 8-ke instead of a 4-ke frequency for the digital discriminator, doubling the effective sensitivity.

V. PILOT PHASE-LOCK

The characteristics of the pilot phase-lock circuit are given in conventional automatic control system terms by Fig. 8. The digital discriminator and filter perform the phase error detection and develop an amplified dc voltage proportional to the phase error. This dc voltage is applied across a varactor diode, causing its capacitance to change and, in turn, produce a change of timing oscillator frequency. The maximum rate of oscillator frequency change is 25 cycles per second. A dc voltage change of 0.15 volt across the varactor diode will cause the oscillator to change frequency one cycle per second.

The response of the digital frequency divider is enough faster than that of the crystal controlled oscillator that its characteristic may be omitted from the control circuit analysis. The sensitivity of the digital discriminator has been made as high as possible to eliminate the need for an additional dc amplifier between the filter and timing oscillator. The steady state error is less than 0.15 microsecond for 64-ke pilot operation and less than 1.2 microseconds for 308-ke pilot operation. The corresponding carrier phase error is dependent on the frequency of the carrier. A maximum time of 0.8 second is required for the timing oscilla-

tor to phase-lock with the pilot frequency. If the pilot is interrupted, the discriminator will generate an average dc voltage of 4.0 volts, corresponding to zero phase error and the timing oscillator will generate a frequency very close to its last calibrated free-running frequency.

The circuit of the digital discriminator, filter, and the varactor portion of the timing oscillator is shown in Fig. 9. The digital discriminator is very similar to the binary counter stage described in Fig. 5. An additional "set" circuit for the pilot is provided at the base of Q_2 and a metering relay is connected between the collectors of Q_1 and Q_2 . When the symmetry of the square waves at the collectors of Q_1 and Q_2 deviates beyond a prescribed amount, corresponding to an oscillator frequency change of one part in 10^6 , the meter relay provides an alarm signal.

The digital discriminator functions in the following manner. Assume the positive going edge of the pilot wave form (see Fig. 9) occurs at t_0 ; the pulse received from the divide-by-256 counter occurs at t_1 , $x_1 = x_2$,

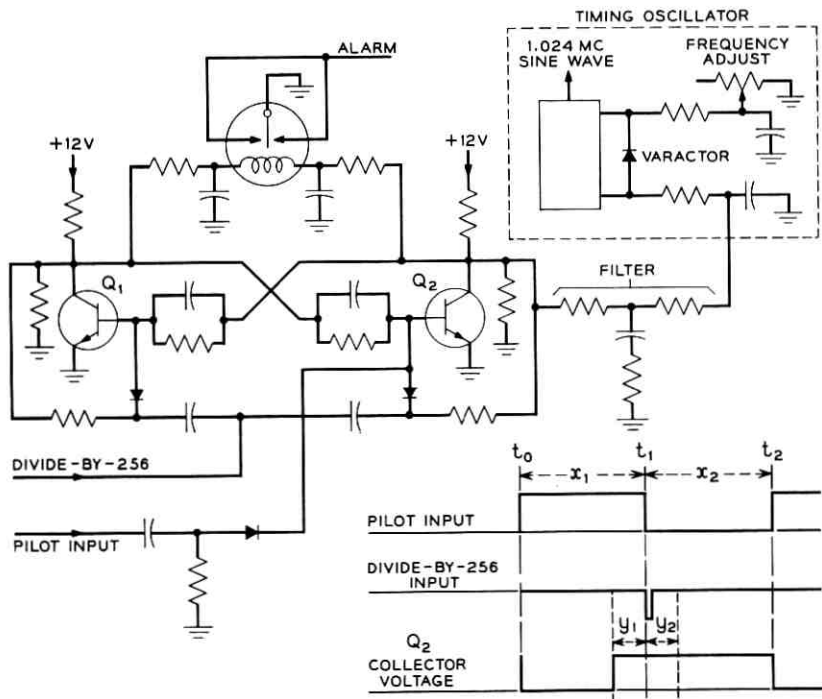


Fig. 9 — Digital discriminator and filter circuits.

and $y_1 = y_2 = 0$. In this condition the output of the discriminator is a symmetrical square wave developing an average of 4.0 volts, the potential used to calibrate the timing oscillator. If the pulse from the divide-by-256 unit arrives early, Q_2 is turned off early, resulting in an output voltage higher than 4 volts. The varactor circuit is arranged so that a higher discriminator voltage results in a lower voltage across the varactor. This lower varactor voltage results in an increased capacitance to the oscillator circuit, causing it to run slower. As it runs slower, the time y_1 , approaches zero, and phase-lock is obtained. A similar set of events occurs if the pulse from the divide-by-256 arrives late with y_2 approaching zero.

VI. AUTOMATIC GENERATOR SWITCHING AND ALARM UNITS

Two separate frequency generators, with automatic switching to either, are provided in the primary frequency supply, as shown in Fig. 10. Either generator A or generator B may be selected manually to permit maintenance on the idle generator or selected automatically in case

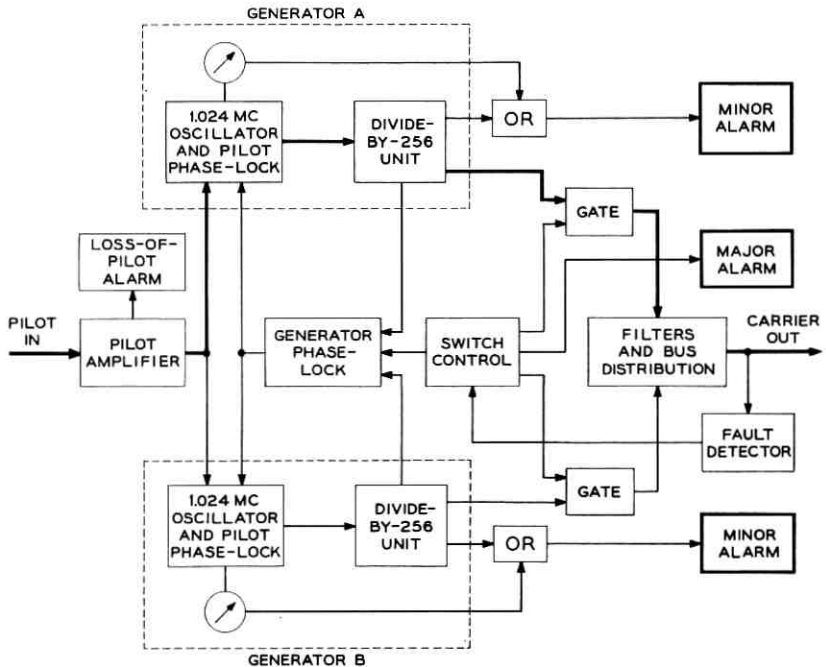


Fig. 10 — Generator switching and alarm units.

the carrier signal disappears at the distribution bus. The heavy lines in Fig. 10 show the signal path through generator A with a similar path available through generator B. Only one pilot amplifier with its own alarm is provided for the two generators, since the timing generator will hold adequate frequency for a short period of time. Two minor alarm circuits are provided with each generator. A meter-relay unit (Sensitrol) is used to monitor the oscillator frequency and a logic circuit is used to detect the operation of the divide-by-256 unit. If a fault is detected by either of these alarm circuits, a minor alarm is initiated but no action is taken by the automatic switching circuit. A fault detector circuit at the distribution bus monitors the presence of the 128-ke, 64-ke, and 4-ke signals. If any of these three signals disappears for longer than one millisecond, the switch control circuit will automatically effect a transfer to the idle generator. If the signals do not reappear at the fault detector within 0.1 second after a transfer, a major alarm is initiated. Transmission "hits" are minimized during manual transfer of generators by operating the two generators phase-locked to each other. A generator phase-lock circuit that detects which generator is in use and phase-locks the idle generator to the active generator is incorporated as part of the logic switching circuits.

VII. OUTPUT CIRCUITS

The outputs of the 4-ke AND gates are connected to a hybrid coil such that either can pass signals to the load and transmission from one side is not effected by either an open or a short circuit steady state condition of the gating transistor on the other side. Thus, 4 kc is supplied to the distribution bus despite failure of one side.

Combining of the 64 and the 128-ke outputs is done in the same way. The square-wave outputs of the combining circuits are reduced to sine waves by filters to insure compatibility with both old and new harmonic generating equipment to which these signals are supplied.

VIII. RELIABILITY

A calculation of the reliability of the primary frequency supply has been made, using failure rate estimates for the various components. Though this calculation is necessarily approximate in the absence of statistical data on the components, it has been useful in striking a balance between the common and redundant sections of the equipment. It also has been influential in determining the allocation of spare plug-in units.

The summation of component failure rate weightings for each section of the equipment results in a "mean time between replacements" for the redundant portion of the supply — oscillator counters and most of the alarm and control circuits — as well as the common equipment. These quantities were 23,000 hours and 250,000 hours, respectively.

Failure of one side of the supply, of course, does not result in interruption of any of the output signals except in the rare case where the other side had failed previously and had not yet been repaired. Thus, in order to combine these statistics and obtain a single statement of the reliability of the supply it is necessary to consider maintenance. Suppose, for example, that spare packages for the unit are available in each office so that only 1 hour is required to repair a defective oscillator. The mean time between coincident failures of the redundant sections is the square of the mean time between replacements ($23,000^2$) divided by the repair time or 530×10^6 hours. If spares are stocked in regional centers so that a week is required, a coincident failure would occur once in 3×10^6 hours on the average. Even under these conditions the reliability of the primary frequency supply is controlled by the common equipment and hence is equal to about 230,000 hours or 26 years.

The mean time between coincident failures is materially reduced when longer repair times are assumed. Thus, the plan for providing spares is based on keeping this interval less than a week.

It is of interest to note that had the primary frequency supply been designed without redundant sides the mean time between failures would have been approximately 3 years. Hence, an improvement of nearly 10 to 1 has resulted from the use of redundancy.

IX. EQUIPMENT FEATURES

Reliability has played a key part in the choice of equipment features. It has dictated that (1) certain functional units be plug-in for ease and expedition of replacement and to facilitate repair at maintenance centers, (2) special care be taken to minimize the operating temperature of the components. The number of circuits that had to be accessible as plug-in units made it desirable that a drawer structure be used. The use of a drawer also made it possible to provide thermal decoupling between the power regulators and the remainder of the components; further, it has minimized the chance of inadvertent removal of the plug-in units, since they are not accessible with the drawer in place.

Figs. 11 and 12 show the complete unit. In the area beneath the

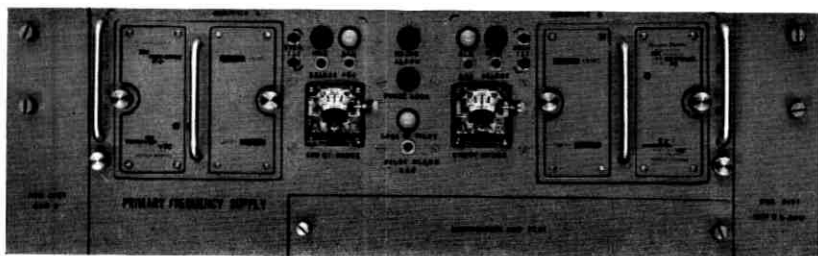


Fig. 11 — Front panel showing alarm and phase indicators and controls.

drawer are a terminal block for external connections and the bus equipment and the jacks. Power regulation units are outside the box. The unit is shown equipped for 308-ke synchronization.

Figs. 13 and 14 show the counter plug-in units. Each three-transistor binary cell occupies a $1\frac{1}{2} \times 5$ inch printed wiring board. Six of these are mounted in the frame of the unit; and soldered interconnections are made at the edges of these to a 4×5 inch board. The associated

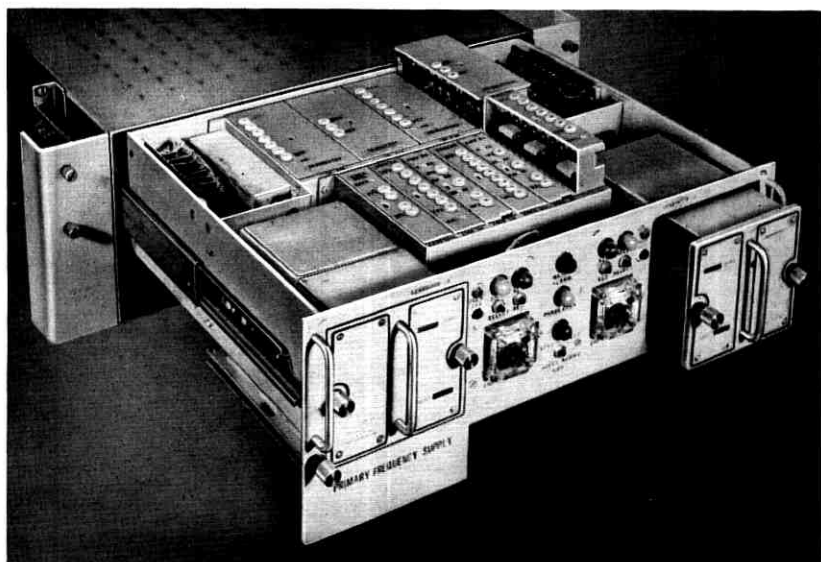


FIG. 12 — Drawer-type construction makes plug-in units accessible but conserves front panel space.

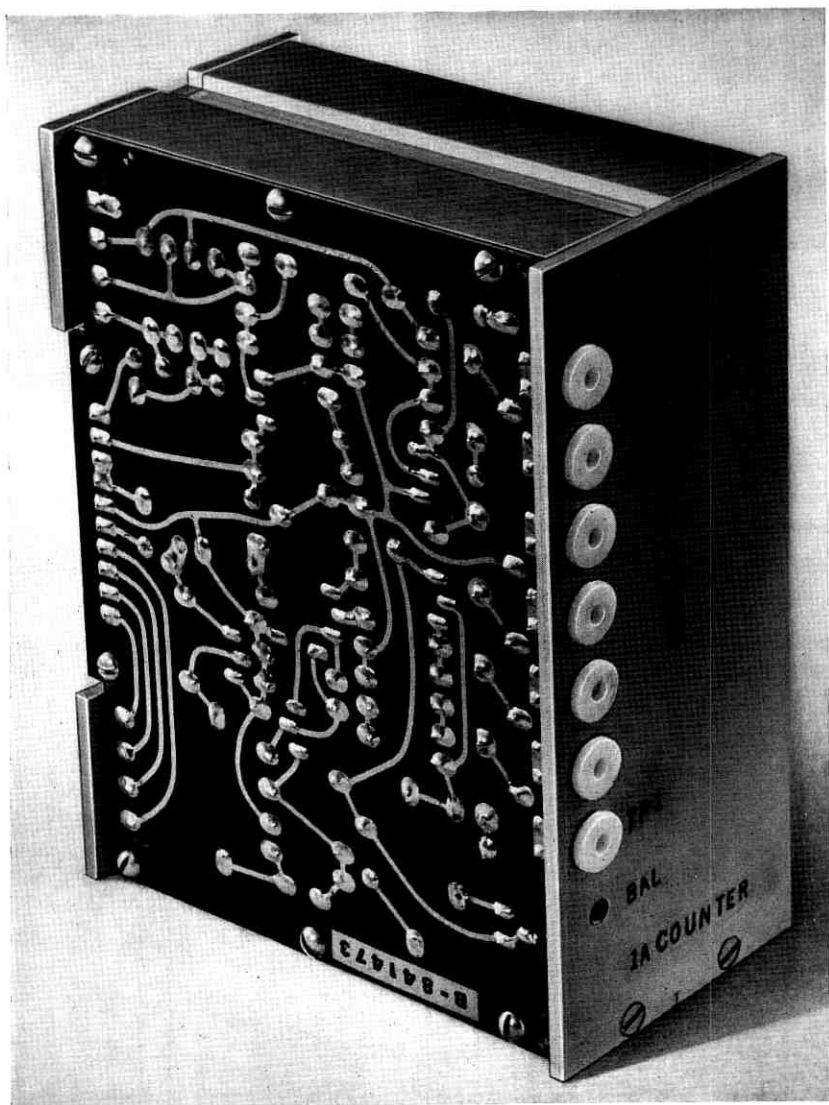


Fig. 13 — Keyway in edge of plug-in unit, location of plug and shape of unit insure proper positioning and orientation.

gate and reset circuits are included on another 4×5 inch board. When these eight boards are fastened in the frame no components are visible on the outside of the resulting package.

The various plug-in units are "keyed" so as to prevent their insertion

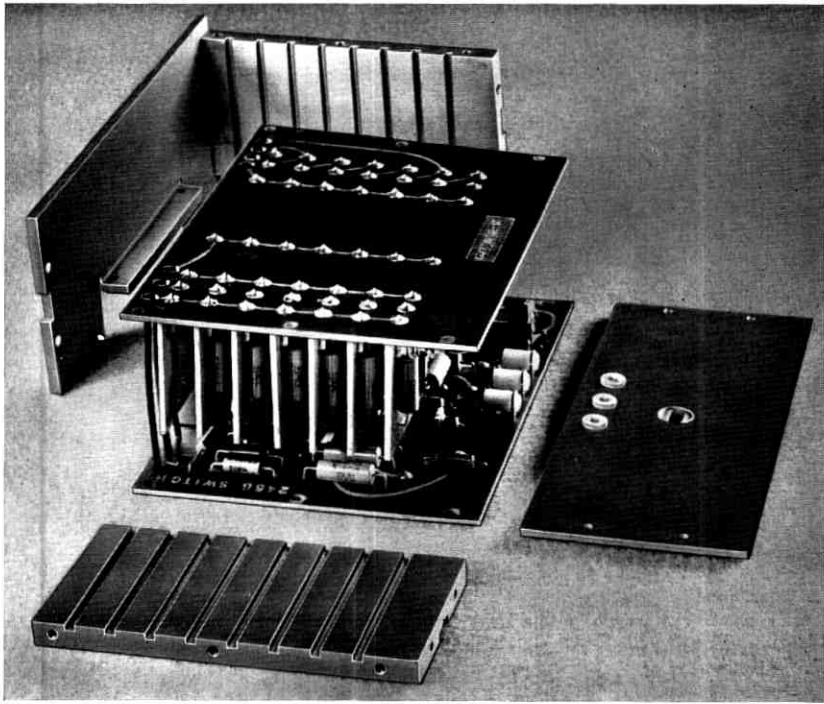


Fig. 14 — Counter units use sandwich-type construction.

in the wrong locations or orientations. Typical units are shown in Figs. 15 and 16.

X. CONCLUSION

Calculations, verified by tests, have indicated that a unit-to-unit delay range equal to 0.08 microsecond when synchronized to 64 kc, and 0.18 microsecond when synchronized to 308 kc, may be expected among accurately adjusted supplies. Any one unit is expected to display a delay variation less than 0.05 microsecond over a 24-hour period, and a gradual change due to crystal aging, not in excess of 0.3 microsecond for 64-kc synchronization or 2 microseconds for 308-kc synchronization. The total phase difference, at 64 kc, between the two sides of a supply is therefore less than 10 degrees.

The maximum frequency error that would result on loss of synchronization is not expected to exceed one part in 10^7 . Routine adjustment of the oscillators not more frequently than once per month is expected to guarantee this accuracy.

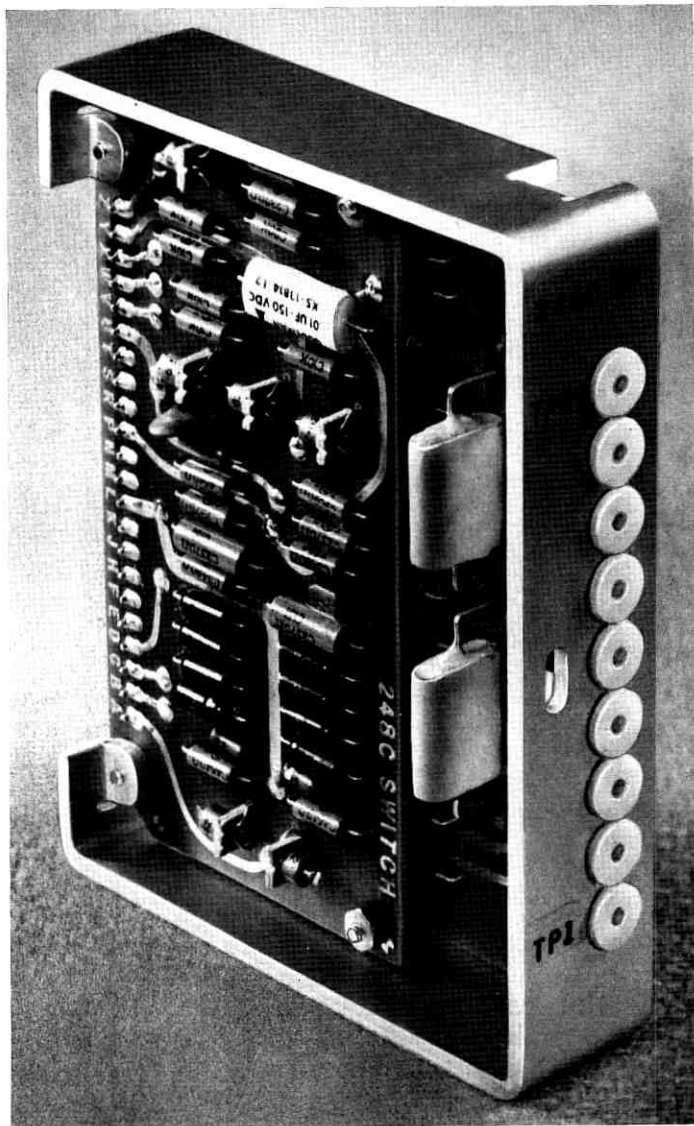


Fig. 15 — This unit contains two printed-wiring boards interconnected at the plug.

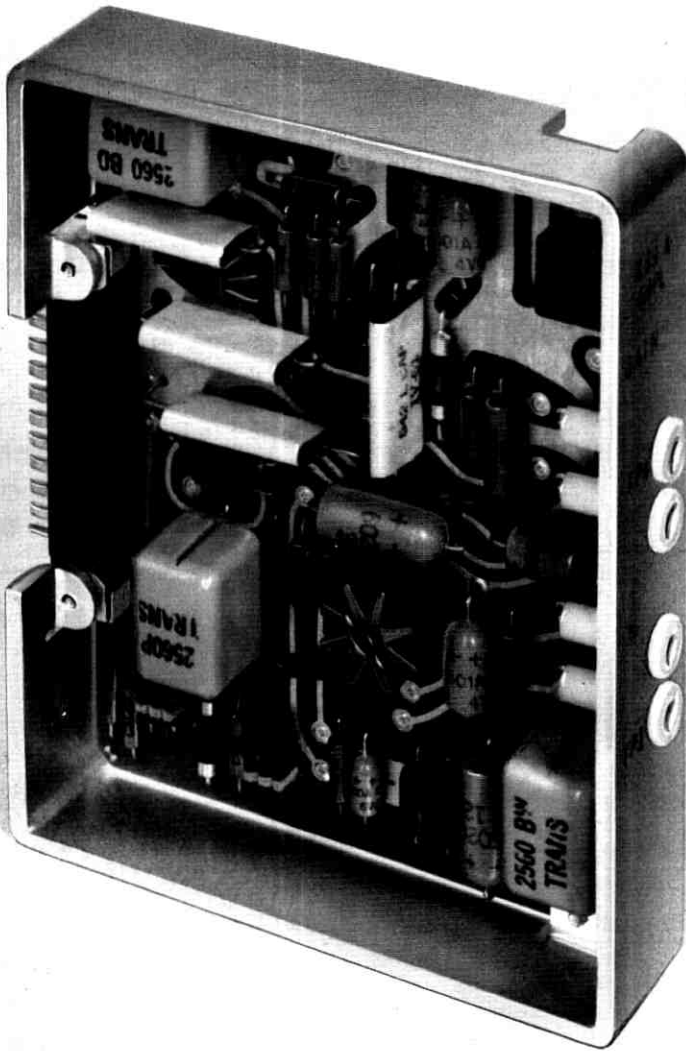


Fig. 16 — Each plug-in unit occupies an individual cell in the drawer. This affords good shielding between units so that individual covers are not required.

In addition to this indicated frequency and phase stability, the requirements outlined in Section II as to compatibility with old and new L multiplex and reliability have been met with margin. An over-all size

reduction of more than 6 to 1, compared with the equivalent units of the old system, was also achieved.

XI. ACKNOWLEDGMENTS

The authors wish to acknowledge the contributions of many others in this development effort, particularly those of Mr. L. F. Travis, who was responsible for the mechanical aspects of the primary frequency supply and Messrs. W. L. Smith and H. S. Pustarfi, Jr., who developed the oscillator unit, as well as Messrs. F. J. Hallenbeck and J. J. Mahoney, Jr., for their many helpful suggestions.

REFERENCES

1. Albert, W. G., Evans, J. B., Ginty, J. J., and Harley, J. B., Carrier Supplies for L-Type Multiplex, B.S.T.J., this issue, p. 279.
2. Crane, R. E., Dixon, J. T., and Huber, G. H., Frequency Division Techniques for a Coaxial Cable Network, Trans. A.I.E.E., **66**, 1947, p. 1451.
3. Line Transmission Maintenance, XVII Plenary Assembly, International Telephone Consultative Committee, Vol. III, Oct., 1954, p. 45.
4. Graham, R. S., Adams, W. E., Powers, R. E., and Bies, F. R., New Group and Supergroup Terminals for L Multiplex, B.S.T.J., this issue, p. 223.
5. Smith, W. L., Miniature Transistorized Crystal — Controlled Oscillator, I.R.E. Trans. on Instrumentation, Vol. **I-9**, No. 2, September, 1960.
6. Byrne, C. J., Properties and Design of a Phase-Controlled Oscillator with a Sawtooth Comparator, B.S.T.J., **41**, March, 1962, p. 559.

Telephone Switching and the Early Bell Laboratories Computers

By E. G. ANDREWS

(Manuscript received October 11, 1962)

This article summarizes the salient features of seven relay-type computers designed and built at Bell Telephone Laboratories. Emphasized are features derived from telephone technology and the role played by the Bell Laboratories designs in advancing the computer art.

I. INTRODUCTION

Various papers, some of them listed in the references, have presented material on two interrelated subjects: (1) automatic data processing as applied to machine switching in the Bell System, and (2) work on digital computation as it grew out of telephone technology and as it was applied to seven relay-type digital computers designed and built at Bell Telephone Laboratories between 1939 and 1950. Together, these papers tell an important story of Bell System contributions to the field of automatic digital computation. It has been felt for some time, however, that a single account of the pertinent facts and developments would be of interest and value. Thus, this paper makes no claims of presenting new material; rather, it is an attempt to bring together many diverse accomplishments and to show their relationships so that a single paper will be available for convenient reference.

II. THE PANEL SYSTEM

Bell Telephone Laboratories and its pre-1925 predecessors, the Western Electric Engineering Department and various American Telephone and Telegraph Company groups, have been engaged with problems of automatic data processing for almost 60 years. The basic reason, of course, is that any automatic telephone switching system must process digital data, beginning with the signals and dialed digits originated by a customer in placing a call.

In the period from about 1903 to 1916, several types of automatic dial

systems were developed, built, and soon abandoned. Starting about 1914, however, two successful systems incorporating central control features were developed: the panel machine switching system¹ used in the United States and the power-driven rotary system used in several European countries. These systems were probably the first in which the familiar electromechanical relay assumed a dominant role in the design. More specifically, the relay opened new possibilities of automatic control. As a consequence, many of the principles of modern computer design were incorporated into telephone switching at a very early period.

The relay, for example, made possible the efficient conversion from one numbering system to another. This was important because it was realized early that the decimal system of notation was wasteful of equipment in the machine handling of numbers. One of these conversions was incorporated into the relay call indicator, subsequently known as the panel call indicator,² used in transmitting called numbers from panel offices and displaying them on lamps at a manual office. At the panel switching office, each decimal digit was converted to a four-pulse code, each pulse having one of two possible signal levels. At the manual office, this special panel call indicator code was converted back to decimal form for display to operators. In another conversion used in panel switching, the first two digits of a customer's telephone number were translated into three digits of an entirely different numbering base, namely 5, 4, 5. (That is, the first digit ranged from 0 to 4, the second from 0 to 3 and the third from 0 to 4.)

The relay also eased the problem of circuit logic. It became easier, for instance, to instrument a system to cope with the all-trunks-busy condition. In other words, upon receipt of an overflow signal, the panel system would automatically change its switching program. This kind of operation is popularly called conditional transfer in modern computers.

Panel switching also provided storage or memory facilities on both a high-speed and low-speed basis. Revertive pulses* were stored in high-speed relay registers; office-code routing information (two-digit codes) was wired in and made available by a slower-speed rotary selector. Both serial and parallel types of operation were employed: serial operation in such circuits as those for revertive pulse counting and call indication, and parallel operation, for example, in reading information on routing.

* In the revertive pulsing arrangement, a switch control sends a start signal to a distant selector, which then begins to advance blindly. Each time the selector advances to a new position, it sends back a pulse (hence the term revertive) to the switch control. When the selector reaches the desired position, the switch control sends a stop signal. A revertive pulse counting circuit was invented by E. C. Molina in 1911. This circuit was an improvement on an earlier counting circuit used in a rotary system developed by F. R. McBerty.

In the 1920's, the valuable bistable flip-flop circuit³ came into use, called the "W and Z" relay combination in telephone switching. In addition, the decoder type of office-code translator⁴ used the direct-access type of table look-up, replacing in some applications the earlier cyclic look-up operation. Also, Boolean algebra for optimizing logic circuitry was being studied, beginning in the late 1930's, by G. R. Stibitz, C. E. Shannon, W. Keister, and others.

In brief, then, many of the more important computer operations and concepts were, in the 1920's and 1930's, already incorporated into the panel switching system, and, as the value of the new circuits was recognized, into other systems as well. It might be asked, therefore, why automatic digital computers were not developed simultaneously. In the period from 1925, for example, this writer heard many of his colleagues say with firm conviction that relay circuit techniques could be used to carry out arithmetical operations commonly performed with desk-type adding machines. The reasons are the traditional ones of economics and failure to see a compelling need. But more importantly, no one had taken the initiative to make a thorough study of all aspects of such a project.

III. THE COMPLEX NUMBER COMPUTER (MODEL I)

In 1937 G. R. Stibitz, who had been employed at Bell Laboratories as a research mathematician since 1930, started to make some significant observations. He noted that Bell Laboratories employed a staff of girls using hand-operated calculating machines for solving the many types of filter network design problems resulting from the introduction of major improvements in voice transmission systems. He found upon closer examination that the bulk of the computing involved addition, subtraction, multiplication and division of complex numbers. Stibitz thereby established one of the prerequisites for a new computer development, namely, he saw a need for it.

Stibitz also knew that another area at Bell Laboratories had for many years been working on dial switching systems as described above. He noted that many of the design techniques therein employed were directly applicable to the design of an automatic computer. He also noted that, for computation, decimal codes would be even less desirable than for switching. Thus, he initially decided in favor of the true binary code for representing numbers. This was a new and revolutionary idea in the computing field, even though small numbers had previously been used in binary form in the panel system decoder and in some European systems. Having made this fundamental decision, he then proceeded to design a relay-type computer. Stibitz thereby satisfied another pre-

requisite for a new computer development, namely, he demonstrated that it was a feasible undertaking.

Briefly, in the original paper design this computer was pictured as having a control board which accepted information from a bank of keys and which displayed output information — that is, the answers — on a bank of switchboard-type lamps. It handled seven-digit decimal numbers with the decimal point at any desired position. The arithmetic unit used an ingenious scheme for making the decimal-to-binary and binary-to-decimal conversions. The devices used were primarily telephone-type relays and slightly modified Strowger-type step-by-step switches.

Stibitz's design was a convincing demonstration that the development of such a computer should be undertaken and that it was an economically feasible project. It was decided that the scope of this computer should initially be limited to handling the four complex number computing operations. This limitation had the two-fold advantage of testing out the principles of an automatic computer on a small but adequate scale and of providing a machine that would perform useful services for the Laboratories. From this time, the computer gradually assumed the name of the "Complex Number Computer," generally shortened to "Complex Computer."

Stibitz studied his design intently. He incorporated many improvements, two of which led to marked simplification. One of these was that he abandoned the provision for a variably located decimal point. All real numbers, therefore, were converted by the human operator to numbers between $+1$ and -1 . It would therefore be necessary for the operators to keep in mind the various scale factors and apply the proper resulting scale factor to the answer. The other major simplification was directed toward reducing the complexity of the necessary binary and decimal conversion facilities. The solution was a hybrid type of code for number representation now commonly used in modern computers and called the "plus-three binary-coded decimal" or BCD. Each decimal 0 to 9 was represented by binary numbers 0011 to 1100. This code provided most of the calculating simplicity of the pure binary code and at the same time preserved the decimal form, so that the thousands, hundreds, etc. digits were always recognizable. Furthermore, the code facilitated the inter-decimal digit carry operations, the complementing operation, and the algebraic sign determination. The plus-three system significantly lowered the cost of the arithmetic unit.

In November, 1938, the responsibility for development of the Complex Computer was placed in the hands of Samuel B. Williams, whose career had for many years been on the front lines of new telephone

systems developments. As a consequence of his study of the project, Williams showed that improved operating conditions could be obtained by using push-button keys for the input and commercially available teletypewriter equipment for the output. The wisdom of this choice was to become apparent later. In the interest of simplifying the circuitry, Williams also chose to use the new crossbar switch instead of the step-by-step switch. The Complex Computer was the only one of the seven relay computers built at Bell Laboratories to employ crossbar switches.

Of course, there were numerous smaller changes in design arrived at as a result of frequent exchanges of ideas between Stibitz and Williams. Perhaps the most important of these was the decision to use two identical calculator units, one for handling the real terms of the complex numbers, and the other for handling the imaginary terms. This arrangement halved the over-all computing time with little extra development effort.

By March, 1939, the equipment design, under the direction of C. E. Boman, was well under way, and in the following month construction was started under the direction of A. J. Bendernagel. Construction was completed by October, 1939. The test period which followed indicated the need of the usual minor changes; these were soon completed, and the computer was moved to its permanent location at the New York City Bell Laboratories location at 463 West Street. The dreams of Stibitz and Williams had now been realized. The Complex Computer was placed in routine operation on January 8, 1940.

The Complex Computer as finally built is illustrated in the two accompanying photographs. Fig. 1 shows one of the three operator stations. These were located in three separated areas in the West Street building. Thus, the first Bell Laboratories relay computer included a feature that is desirable in many modern installations: multiple input positions with the lockout facilities required for this type of operation.

Fig. 2 shows the computing equipment. The computer was an eight-place decimal machine with two extra digits provided in the arithmetic unit to compensate for excessive round-off errors when accumulating many subtabulations. While this Complex Computer was specifically designed to handle the four complex number arithmetic operations, the users soon found that, with certain special variations, the computer could be used advantageously for other types of problems.

IV. THE HANOVER DEMONSTRATION

But Stibitz and Williams had still more planning and designing to do. It had been decided to demonstrate the principles of the computer to the



Fig. 1 — One of the three operator stations of the Complex Computer.

public. Due to the nature of the device, it was further decided that the demonstration would be most effective if carried out before a mathematically minded audience. The scene chosen, therefore, was Hanover, New Hampshire, on the occasion of the fall, 1940, meeting of the American Mathematical Society. Williams designed new equipment to provide a station at Hanover and new facilities for operating the computing equipment remotely over a teletypewriter circuit.⁵ The demonstration was held at the meeting of September 11, 1940. Stibitz read a paper describing the machine, after which he and T. C. Fry showed how problems could be introduced at the scene of the conference and transmitted to the Complex Computer in New York. They were then able to point to the teletypewriter, which typed out the answer within a minute.⁶ Those interested were invited to introduce their own problems and check the results. The demonstration was also significant because, in employing a

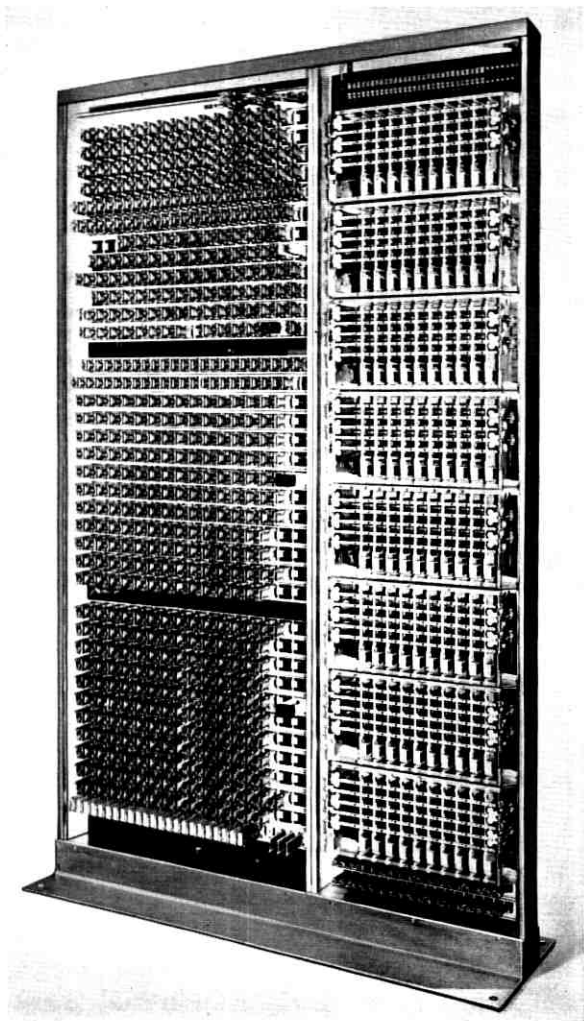


Fig. 2—The computing equipment of the Complex Computer.

data link between Hanover and New York, it foreshadowed the present growth of data transmission services. During this demonstration, Williams monitored the computing equipment in New York, but it was wasted effort. The system worked flawlessly.

With the advent of World War II, demands on the Complex Computer increased, and many military projects moved a little faster toward com-

pletion because of its contribution. As other computers were built, the Complex Computer came to be called Model I. It was in daily use at Bell Laboratories until 1949.

Table I lists many of the logical and physical design features of Model I and of the five later models of computers. Two units of Model V were built, so that a total of seven relay-type computers were built at Bell Laboratories. Also, Model V could be expanded to a system of six computers and ten problem positions, although the two units were built with only two computers each and with either three or four problem positions.

Before discussing Models II through VI, however, it is well to gain some perspective of the course of computer design at Bell Laboratories. Model I was a real advance in applying electromagnetic techniques to machine calculation, but of course by today's standards it would seem to be little more than an elaborate desk calculator. Model II was improved in many respects, but was essentially designed to perform only one type of calculation. Models III and IV represented more significant advances in that they were the first Bell Laboratories machines that in today's sense could be called programmable, but only for restricted classes of problems. Models V and VI, by contrast, were truly modern computers. Although slower than the more recent electronic computers, they were organized much like present machines. All five models, however, steadily advanced the computer art and anticipated many features that even today are just beginning to be widely used.

V. MODELS II-IV

Model II^{7,8,9} was built for the National Defense Research Council, and was placed in operation in September, 1943. A truly special-purpose computer, Model II was designed to handle certain fire-control problems that took only a few months to solve. Afterwards, however, many other tasks were found for it. Model II used the bi-quinary type of number representation with its self-checking features. It used a paper tape input and output; the program was perforated on a tape formed into a loop. The control unit recognized 31 instructions, which included the addresses of the registers. Model II was retired from service by the U. S. Naval Research Laboratory in early 1961.

Model III,¹⁰ the basic development of which was completed in June, 1944, was another computer built for the National Defense Research Council and used for military fire-control problems. However, it handled many types of problems creditably, and was to some extent a general-purpose machine. Many features of Model III were new, of chief im-

TABLE I — STATISTICAL INFORMATION ABOUT BELL LABORATORIES COMPUTERS, MODELS I TO VI

	Model I	Model II	Model III	Model III ^c	Model IV	Model V	Model VI
LOGICAL DESIGN FEATURES							
Number of built-in routines.....	2	0	0	0	0	4	200
Decimal point.....	Fixed	Fixed	Fixed	Fixed	Fixed	Floating	Floating
Discriminating action.....	None	Note 1	Note 1	Note 1	Note 1	Exten.	Yes
Multiplication.....	Yes	Note 2	Yes	Yes	Yes	Yes	Yes
Division.....	Yes	No	No	No	No	Yes	Yes
Square Root.....	Yes	No	No	No	No	Yes	Yes
Indeterminate arithmetic.....	No	No	No	No	No	Yes	Yes
Special trigonometric features.....	No	No	Note 1	Note 1	Note 1	Yes	No
Special logarithmic features.....	No	No	No	No	No	Yes	No
Round off — automatic or program.....	No	Pro.	Pro.	Pro.	Pro.	Auto.	Auto.
Subscript knowledge.....	No	No	Yes	Yes	Yes	No	No
Number of addresses in code.....	No	1	1 or 2	1 or 2	1 or 2	3	3
Self checking.....	No	90%	100%	99%	100%	100%	100%
PHYSICAL DESIGN FEATURES							
Number of relays.....	450	440	1,400	≈1750	1,425	9,000	4,600
Pieces of teletypewriter equipment.....	4	5	7	≈10	7	55	16
Number of number registers.....	4	7	10	14	10	15	12
Number of digits per number.....	8	2 to 5	1 to 6	1 to 6	1 to 6	1 to 7	3, 6, 10
Multiplication time in sec. per 5 digit number.....		≈4.3	1	1	1	0.8	0.8
Number of problem stations.....	3	1	1	1	1	3 and 4	2
Arranged for unattended operation.....	No	No	Yes	Yes	Yes	Yes	Yes
Number notation with self-checking binary.....	No	Yes	Yes	Yes	Yes	Yes	Yes
“2 out of 5”.....	No	No	Yes	Yes	Yes	No	No
“3 out of 5”.....	No	Yes	Yes	Yes	Yes	Yes	Yes

^c This column applies to the Model III after its modification in 1949.

Note 1. Very limited application.

Note 2. With multiplier specified in program.

portance being its 100 per cent self-checking of all operations. The machine stopped positively on any kind of single failure and on most combinations of two or more simultaneous failures.

A second and very significant feature of Model III was its calculator, designed by E. L. Vibbard. This unit multiplied and divided by the principle of finding partial products in a multiplication table. Other noteworthy features were table-hunting, double-entry interpolation, and unattended operation. Also included was the feature now commonly called indexing. In the Model III it was termed subscript notation, although it did not have the full flexibility of modern computer indexers. In all, Model III was possibly the most interesting of the Bell Laboratories computers from the point of view of its design logic and of the ease of understanding its operations. In 1944 Model III was moved to Fort Bliss, Texas, and was in use until 1958.

Model IV¹¹ looked like Model III and did the same kind of computing. However, changes were included enabling Model IV to handle trigonometric functions from -90° to $+360^\circ$. It was built for Naval Ordnance and its basic development was completed in March 1945. In Naval circles it was known as the Mark 22, and was in service until early 1961.

VI. MODELS V AND VI

As mentioned above, Models I through IV can be regarded as belonging to an era of the past. Models V and VI, however, bridged the gap between the beginnings of the art and the modern era of electronic computers.

In size and flexibility, Model V¹²⁻¹⁶ was Bell Laboratories' most ambitious project in computer development up to that time. Two were built. One was delivered (1946) to the National Advisory Committee on Aeronautics, and the second was built (1947) for the Ballistics Research Laboratory at Aberdeen, Maryland.

As mentioned earlier, Model V was a system having a maximum of six computers and ten problem positions. Such arrangement permitted the computers to function continuously. Problems were loaded into idle problem positions, and a computer on completion of one problem automatically picked up another. A problem position had one tape reader for input data, up to five readers for the program of instructions, which allowed considerable flexibility in introducing sub-routines, and up to six readers for tabular data. Tables of logarithms, antilogarithms, sines, cosines and antitangents were permanently wired into the machine.

The calculator of Model V included: floating decimal point, multiplica-

tion by "short-cut" addition, automatic roundoff (but subject to cancellation), the ability to recognize most indeterminate arithmetic operations, special facilities for trigonometric and logarithmic calculations, and special auxiliary equipment for processing of various paper tapes. Model V also included rather elaborate discriminatory controls. These controls, as in the panel switching system application mentioned earlier, are now referred to as conditional transfers. Model V operated around the clock, and had an excellent record for low out-of-service time.

In recent years, the two Model V computers have had an interesting history. The unit at Aberdeen was transferred to Fort Bliss and later was given to the University of Arizona for educational and research programs. The unit built for the National Advisory Committee on Aeronautics was given to Texas Technological College in early 1958. Unfortunately, however, it was severely damaged in transit and was of no further use except for spare parts for the University of Arizona machine.

Model VI,^{17,18} after an extensive test period, was placed in regular service in 1950. It was built for Bell Laboratories' own use in solving a wide variety of research and development problems. In essence, Model VI was a simplified version of Model V, having only one computer and less elaborate discriminatory controls and problem positions. In other respects, however, Model VI incorporated many improvements. It had three storage tapes, one of which could have either numbers or instructions. It had a system of several hundred semipermanent subroutines, each of which could be used for problems of the same type with differing numbers of variables. The internal subroutine control adjusted itself to the particular problem at hand. Model VI also had what is called the "end of numbers" check signal, which eased such problems as determining the end of a line of coefficients in a matrix-type problem. Also, Model VI had an automatic "second trial" feature that functioned during unattended operation, thus improving reliability in the presence of a trouble condition. After several years of service at Bell Laboratories, Model VI was given to the Polytechnic Institute of Brooklyn. In 1960, Brooklyn Poly retired this computer and gave it to the Bihar Institute of Technology in Bihar, India.

VII. CONCLUSION

With this background of Bell Laboratories work in the computer field, and of the telephone technology that prompted the developments, it is possible to make a few comments concerning the extent of the total contribution. Aside from their specifically mathematical aspects, these con-

tributions are chiefly in the areas of machine accuracy and dependability, and ease of programming, operation and maintenance.

On the score of accuracy, through 1951 only two errors were reported as resulting from machine faults in Models III through VI. This enviable record reflects the corresponding accuracy of telephone switching systems and resulted partly from such telephone-derived features as self-checking and second-trial functions. The dependability of the computers also derived in large measure from the extensive heritage of experience in designing switching systems. In particular, notice must be taken of the U-type relay and its associated circuits. After many years of careful design, the U-type relay had become a symbol of the reliability of telephone systems, and it was used to great advantage in all the Bell Laboratories relay computers.

Model V compared favorably with other contemporary computers in ease of programming, and Model VI brought about an even greater reduction of effort. With problems using repetitive subroutines, programming effort with Model VI was near the minimum. In terms of operating ease, no more than five minutes was required to load Model V with the most complex problem it could handle, and even less time was required with Model VI. Generous use of indicating signals aided in analyzing machine stoppages, so that maintenance was thereby facilitated. For both operation and routine maintenance, highly trained personnel were not required.

These are the same factors stressed in telephone systems design, so the above summary could be an apt description of almost any switching system developed at Bell Laboratories. The similarity again stresses the way in which telephone research and development have had an important influence on computers.

VIII. ACKNOWLEDGMENTS

I am indebted to N. E. Sowers, Scientific Advisor, U. S. Army Air Defense Board, to Dr. E. J. Smith, Professor of Electrical Engineering at Polytechnical Institute of Brooklyn, and to J. J. Fleming, Electronics Scientist, U. S. Naval Research Laboratory, for certain details concerning the post-Laboratories history of the Models II to VI computers.

REFERENCES

1. Craft, E. B., Morehouse, L. F., and Charlesworth, H. P., Machine Switching Telephone System for Large Metropolitan Areas, B.S.T.J., **2**, April, 1923, p. 53.

2. *Ibid.*, p. 80; also see Gherardi, B., and Charlesworth, H. P., Supplement to the Telephone Review, April, 1920; Aitken, W., *Automatic Telephone Systems*, Benn Brothers, Ltd., London, Vol. 1, 1921, p. 180; Kloeffler, R. G., *Telephone Communications Systems*, Macmillan, New York, 1925, p. 211.
3. Patent number 1,751,263, Electrical Switching System, 3/18/30, O. Cesareo.
4. Raymond, R., The Decoder, Bell Laboratories Record, **6**, May, 1928, p. 273; McAlpine, R. K., Automatic Prevention of Trouble by Decoders, Bell Laboratories Record, **8**, July, 1930, p. 518.
5. Patent number 2,434,681, Remotely Controlled Electrical Calculator, 1/20/48, S. B. Williams.
6. Fry, T. C., Amer. Math. Soc. Bulletin, **46**, 1940, p. 861. Also see Sci. News Letter, Sept. 14, 1940; New York Sun, Sept. 9, 1940; Newark News, Sept. 10, 1940; New York Herald Tribune, Sept. 15, 1940.
7. Cesareo, O., The Relay Interpolator, Bell Laboratories Record, **24**, Dec., 1946, p. 457.
8. The Relay Interpolator, Final technical report to the National Defense Research Council, Bell Telephone Laboratories, Oct. 31, 1945.
9. The Relay Interpolator, Report No. R-3177, Naval Research Laboratory (Washington, D. C.), Sept. 25, 1947.
10. Juley, J., The Ballistic Computer, Bell Laboratories Record, **25**, Jan., 1947, p. 5.
11. Computer Mark 22, Model O. Report 178-45, United States Navy (Washington, D. C.), Dec. 12, 1945.
12. Williams, S. B., A Relay Computer for General Application, Bell Laboratories Record, **25**, Feb., 1947, p. 49.
13. Williams, S. B., Annals, Harvard Computation Laboratory (Cambridge, Mass.), **16**, p. 41, 1947.
14. Alt, F. L., Mathematical Tables and Other Aids to Computation, National Research Council, **3**, Pt. I, No. 21, p. 1, and Pt. 2, No. 22, p. 69.
15. Relay Computer for the Army, Bell Laboratories Record, **26**, May, 1948, p. 208.
16. Bode, H. W., and Andrews, E. G., Use of the Relay Digital Computer, Electrical Engineering, **69**, Feb., 1950, p. 158.
17. Andrews, E. G., Annals, Harvard Computation Laboratory (Cambridge, Mass.), **26**, p. 20, 1951.
18. Andrews, E. G., The Bell Computer, Model VI, Electrical Engineering, **68**, Sept. 1949, pp. 751-56.

On the Theory of Linear Multi-Loop Feedback Systems

By I. W. SANDBERG

(Manuscript received October 19, 1962)

It is well known that the concept of return difference plays a central role in the classical theory of linear feedback systems developed by Black, Nyquist, Blackman and Bode. This concept, which relates to the influence of a single algebraic system-constraint of the form $f_2 = \gamma f_1$ where f_1 and f_2 respectively may be regarded as a controlling signal and a controlled signal, retains its prominence in the subsequent signal-flow graph theoretic extensions by Mason. It is particularly pertinent to the study of the stability of the system, its degree of immunity from parameter variations, and the determination of its transmission and driving-point properties.

This paper reports on a generalization of Blackman's equation and on some generalizations of Bode's return difference theorems. Here attention is focused not on a single constraint of the form $f_2 = \gamma f_1$, but instead on a set of constraining equations.

I. INTRODUCTION

It is well known that the concept of return difference plays a central role in the classical theory of linear feedback systems developed by Black,¹ Nyquist,² Blackman³ and Bode.⁴ This concept, which relates to the influence of a single algebraic system-constraint of the form $f_2 = \gamma f_1$ where f_1 and f_2 respectively may be regarded as a controlling signal and a controlled signal, retains its prominence in the subsequent signal-flow graph theoretic extensions by Mason.⁵ It is particularly pertinent to the study of the stability of the system, its degree of immunity from parameter variations, and the determination of its transmission and driving-point properties.

This paper reports on a generalization of Blackman's equation³ and on some generalizations of Bode's return difference theorems.⁴ Here attention is focused not on a single constraint of the form $f_2 = \gamma f_1$, but instead on a set of constraining equations. For this reason the results

are applicable to systems which possess in the usual physical sense a multiplicity of feedback loops. However it will become clear that the results are not restricted to situations of this type.

In Section II we describe the basic system considered throughout the paper. Section III presents an explicit expression for $w(\mathbf{X})$, the transmission or driving-point function to be studied. We then introduce definitions, based on a simple topological characterization of the relation between the system input and output variables, of the loop-difference matrix, the null loop-difference matrix, and the complementary loop-difference matrix. The determinants of these matrices are of fundamental importance in the subsequent discussion. In Section VI we derive a generalization of Blackman's interesting equation. The material in Sections VII and VIII relates to generalizations of Bode's well-known return difference theorems. Some applications and illustrations of the theory are discussed in Section IX.

II. THE BASIC SYSTEM AND THE SET OF EQUATIONS \mathfrak{F}

We shall be concerned throughout with the transmission and driving-point properties of the structure shown in Fig. 1, an arbitrary linear time-invariant two-port network containing no independent sources.

The Laplace-transformed equilibrium equations which implicitly define the external properties of the two-port are of the form:

$$\sum_j (\alpha_{kj} e_j + \beta_{kj} i_j) = 0 \quad (k = 1, 2, \dots, K) \quad (1)$$

where e_j and i_j respectively are branch voltages and currents and the α_{kj} and β_{kj} are functions of the complex-frequency variable.

We wish to focus attention on the influence of a prescribed subset of the linear constraints implied by (1). This subset is assumed to be expressible as

$$f_k = \sum_{i=q+1}^{q+p} \gamma_{ki} f_i \quad (k = 1, 2, \dots, q) \quad (2)$$

in which each f_i is one of the e_j or one of the i_j . It is convenient to interpret these relations as corresponding to a set of q controlled (i.e.,

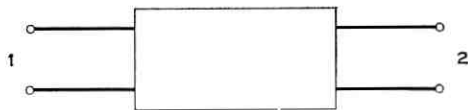


Fig. 1 — Two-port network.

dependent) sources, where the quantity on the left-hand side of each equation is a controlled-source output and each f_i on the right-hand side is a controlling variable. Thus we take the viewpoint that the two-port network contains a controlled-source subnetwork characterized by (2) even when the two-port does not contain devices such as vacuum tubes or transistors with which controlled-sources are ordinarily associated. We shall denote by \mathfrak{F} the set of controlled-source constraints (2).

Let the components of a q -vector Ψ be the f_k ($k = 1, 2, \dots, q$) arranged in any one of the $q!$ possible orders, and let the components of a p -vector Φ be the controlling variables f_i ($i = q + 1, q + 2, \dots, q + p$) arranged in any one of the $p!$ possible orders. Then (2) can be written as $\Psi = \mathbf{X}\Phi$. We shall call the $q \times p$ matrix \mathbf{X} the controlled-source matrix or the matrix of controlled-source coefficients.

III. EVALUATION OF THE TRANSFER AND DRIVING-POINT FUNCTIONS FOR THE TWO-PORT

Our primary concern here is the determination of the influence of \mathfrak{F} (i.e., of the controlled-source subnetwork) on the transfer and driving-point functions for the two-port network.

Let $y_2 = w(\mathbf{X})y_1$ represent the relation to be studied in which y_2 and y_1 respectively are response and excitation functions (each function may be either a voltage or a current). If $w(\mathbf{X})$ is a driving-point imittance, y_2 and y_1 respectively are a current and a voltage or a voltage and a current at the same port.

Consider now an evaluation of $w(\mathbf{X})$. With y_1 and the components of Ψ treated as independent variables, we apply the superposition theorem to obtain†

$$y_2 = dy_1 + \mathbf{B}\Psi \tag{3}$$

where d and \mathbf{B} are defined by the equation. In particular when $\mathbf{X} = \mathbf{0}$, $\Psi = \mathbf{0}$, and $y_2 = dy_1$. That is, $d = w(\mathbf{0})$.

Similarly we can express Φ as

$$\Phi = \mathbf{A}y_1 + \mathbf{C}\Psi \tag{4}$$

where the matrices \mathbf{A} and \mathbf{C} are defined by the equation. After pre-multiplying both sides of (4) by \mathbf{X} and using $\Psi = \mathbf{X}\Phi$, we find that

$$\Psi = [\mathbf{1}_q - \mathbf{XC}]^{-1}\mathbf{XA}y_1 \tag{5}$$

where $\mathbf{1}_q$ is the identity matrix of order q .

† We are assuming that y_2 is uniquely determined by y_1 and Ψ . See the relevant discussion in the next section.

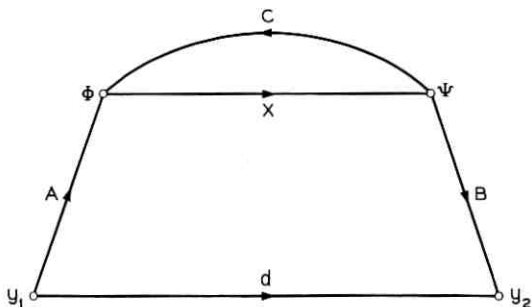


Fig. 2 — Basic flow-graph representation of the relation between y_1 and y_2 .

From (3), (5), and $d = w(\mathbf{0})$

$$w(\mathbf{X}) = y_2(y_1)^{-1} = w(\mathbf{0}) + \mathbf{B}[\mathbf{1}q - \mathbf{X}\mathbf{C}]^{-1}\mathbf{X}\mathbf{A}. \quad (6)$$

IV. THE BASIC FLOW GRAPH

In order to express $w(\mathbf{X})$ in terms of quantities that are related to concepts of importance in the classical theory of linear feedback systems, we first represent the relation between y_1 and y_2 by the signal-flow graph⁵ in Fig. 2. The matrix equation associated with "node-group" Φ is (4), and so forth. This topological characterization of the influence of \mathfrak{F} plays a central role in the subsequent discussion. We shall call \mathbf{A} , \mathbf{B} , \mathbf{C} , and d "flow-matrices."

It is worth emphasizing at the outset that although the signal-flow graph exhibits feedback in the sense that the outputs of the controlled sources (i.e., the elements of Ψ) influence the value of the controlling variables (i.e., the elements of Φ), it is not necessary that feedback exist in the network in the usual physical sense. For example, the physical system may comprise simply a set of driving-point impedances and the controlled sources may represent a certain subset of these one-port elements. Feedback arises in the topological characterization merely because of the form of the equations that we have chosen to write.

Before proceeding it is important to note that for some choices of \mathfrak{F} it may not be possible to characterize the two-port by equations of the form (3) and (4) and hence by the flow graph in Fig. 2. The superposition theorem implies that when y_2 and Φ are uniquely determined by y_1 and Ψ , the relations are of the form (3) and (4); it does not imply that such relations exist. The one-port network shown in Fig. 3 is one

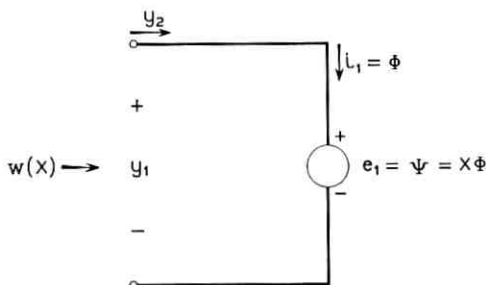


Fig. 3 — Single controlled-source one-port.

of the simplest structures which illustrate the difficulty. For this network, with $w(\mathbf{X})$ the well-defined driving-point admittance, all four flow matrices \mathbf{A} , \mathbf{B} , \mathbf{C} , and d fail to exist.

There are much more sophisticated situations of this general type. Nevertheless, almost all feedback networks of interest do not exhibit this degeneracy. Furthermore when the difficulty does occur, it is generally possible to consider the limiting form of a network for which the degeneracy is not present. For example, if the one-port in Fig. 3 is modified by adding a series resistor, the flow matrices \mathbf{A} , \mathbf{B} , \mathbf{C} , and d become well-defined and the expression obtained for the driving-point admittance reduces to $(\mathbf{X})^{-1}$ as the value of the series resistor approaches zero.

With this motivation we state the

Assumption: It is assumed throughout that the relation between y_1 and y_2 can be represented by a signal-flow graph of the type shown in Fig. 2.

Consider now the definition of three matrices that are related to the flow graph.

4.1 The Loop-Difference Matrix for the Branch \mathbf{X} : $\mathbf{F}_\mathbf{X}(\mathbf{X}_1)$

The loop-difference matrix for the branch \mathbf{X} is defined as follows. Introduce an additional node-group \mathbf{P} in the flow-graph of Fig. 2 by replacing \mathbf{X} with a cascade of any two branches \mathbf{X}_1 and \mathbf{X}_2 such that $\mathbf{X}_1\mathbf{X}_2 = \mathbf{X}$. Let the orders of \mathbf{X}_1 and \mathbf{X}_2 respectively be $q \times m$ and $m \times p$. Next, break the feedback loop by splitting \mathbf{P} into a source node-group \mathbf{P}' and a sink node-group \mathbf{P}'' to obtain the graph in Fig. 4.

With y_1 set equal to zero, suppose that an arbitrary signal vector \mathbf{S}_m , of order m , is applied to \mathbf{P}' . The resulting signals at node-groups

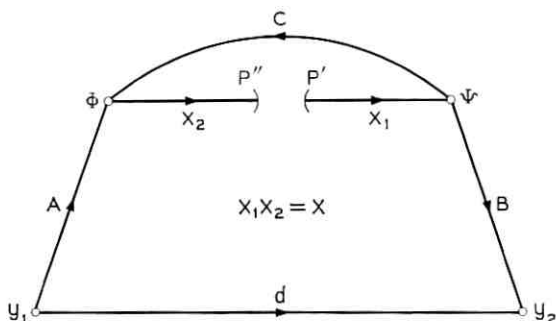


Fig. 4 — Signal-flow graph for defining the loop-difference matrix and the null loop-difference matrix.

Ψ , Φ , and P'' respectively are X_1S_m , CX_1S_m , and $X_2CX_1S_m$. The vector signal difference at P', P'' (i.e., the signal at P' less that at P'') is $[1_m - X_2CX_1]S_m$. We shall call $[1_m - X_2CX_1]$ the loop-difference matrix for the branch X and denote it by $F_{\mathcal{L}}(X_1)$, which indicates explicitly its dependence on X_1 . When X_1 , X_2 , and C are scalars, $F_{\mathcal{L}}(X_1)$ becomes independent of X_1 and reduces to Mason's flow-graph definition of the loop-difference for the branch X .

4.2 The Null Loop-Difference Matrix for the Branch X : $\hat{F}_{\mathcal{L}}(X_1)$

This matrix is evaluated under the condition that y_1 is adjusted so that $y_2 = 0$. Specifically, if an arbitrary signal vector S_m is applied to P' in Fig. 4, the signal reaching y_2 by way of the branch B is BX_1S_m so that, if y_2 is to be zero, y_1 must be $-d^{-1}BX_1S_m$. The total signal arriving at Φ is $-d^{-1}ABX_1S_m + CX_1S_m$, and hence the signal at P'' is $[X_2CX_1 - d^{-1}X_2ABX_1]S_m$. Thus the vector signal difference at P', P'' , under the condition that y_1 is adjusted so that $y_2 = 0$, is $[1_m - X_2CX_1 + d^{-1}X_2ABX_1]S_m$. We call $[1_m - X_2CX_1 + d^{-1}X_2ABX_1]$ the null loop-difference matrix for the branch X and denote it by $\hat{F}_{\mathcal{L}}(X_1)$. When A , B , C , X_1 , and X_2 are all scalars, $\hat{F}_{\mathcal{L}}(X_1)$ reduces to Truxal's definition⁶ of the null loop-difference for the branch X , and is independent of the choice of X_1 . It is convenient to write: $\hat{F}_{\mathcal{L}}(X_1) = [1_m - X_2\hat{C}X_1]$, where $\hat{C} = C - d^{-1}AB$.

4.3 The Complementary Loop-Difference Matrix for the Branch X : $\bar{F}_{\mathcal{L}}(X_1)$

The complementary loop-difference matrix, denoted by $\bar{F}_{\mathcal{L}}(X_1)$, is defined by the requirement that when an arbitrary S_m is applied to

\mathbf{P}' , $\bar{\mathbf{F}}_{\bar{y}}(\mathbf{X}_1)\mathbf{S}_m$ is the vector signal difference at \mathbf{P}' , \mathbf{P}'' under the condition that y_1 is adjusted so that $(y_1 + y_2) = 0$. This matrix is therefore obtained from the expression for the null loop-difference matrix by setting $d = 1$. That is, $\bar{\mathbf{F}}_{\bar{y}}(\mathbf{X}_1) = [\mathbf{1}_m - \mathbf{X}_2\bar{\mathbf{C}}\mathbf{X}_1]$ where $\bar{\mathbf{C}} = \mathbf{C} - \mathbf{A}\mathbf{B}$. The complementary loop-difference matrix is of utility when $d = 0$. In such cases the null loop-difference matrix is not defined.

It is evident that the concepts of loop-difference matrix, null loop-difference matrix, and complementary loop-difference matrix need not be restricted to the branch \mathbf{X} in Fig. 2; they relate without ambiguity to any branch in a signal-flow graph which possesses a single input node and a single output node.

4.4 Circuit Interpretations of \mathbf{C} , $\hat{\mathbf{C}}$, and $\bar{\mathbf{C}}$

It is important to note that the matrices \mathbf{C} , $\hat{\mathbf{C}}$, and $\bar{\mathbf{C}}$ possess very explicit circuit interpretations. When the elements of Ψ (i.e., the outputs of the controlled sources) are treated as independent variables, the p -vector of voltages and currents Φ is equal to $\mathbf{C}\Psi$, $\hat{\mathbf{C}}\Psi$, and $\bar{\mathbf{C}}\Psi$ respectively when y_1 is set equal to zero, y_1 is adjusted so that $y_2 = 0$, and y_1 is adjusted so that $(y_1 + y_2) = 0$. The evaluation of these matrices from the circuit is simplified considerably by the fact that the controlled sources are treated as independent sources.

V. THE DETERMINANTS $\det \mathbf{F}_{\bar{y}}(\mathbf{X}_1)$, $\det \hat{\mathbf{F}}_{\bar{y}}(\mathbf{X}_1)$, AND $\det \bar{\mathbf{F}}_{\bar{y}}(\mathbf{X}_1)$

Our primary interest in the matrices $\mathbf{F}_{\bar{y}}(\mathbf{X}_1)$, $\hat{\mathbf{F}}_{\bar{y}}(\mathbf{X}_1)$, and $\bar{\mathbf{F}}_{\bar{y}}(\mathbf{X}_1)$ is with regard to their determinants. Here we wish to establish two elementary properties of these determinants which add to an understanding of the character of the expression for $w(\mathbf{X})$ presented in the next section. Further properties of these determinants are considered subsequent to the derivation of the expression for $w(\mathbf{X})$.

Although the matrices $\mathbf{F}_{\bar{y}}(\mathbf{X}_1)$ and $\hat{\mathbf{F}}_{\bar{y}}(\mathbf{X}_1)$ generally depend upon the choice of \mathbf{X}_1 , it is true that

Lemma I:

$$\det \mathbf{F}_{\bar{y}}(\mathbf{X}_1) = \det \mathbf{F}_{\bar{y}}(\mathbf{1}_q)$$

$$\det \hat{\mathbf{F}}_{\bar{y}}(\mathbf{X}_1) = \det \hat{\mathbf{F}}_{\bar{y}}(\mathbf{1}_q).$$

The lemma is proved in Appendix A. It implies, of course, that the determinants are independent of the choice of \mathbf{X}_1 . This property is evidently shared by $\det \bar{\mathbf{F}}_{\bar{y}}(\mathbf{X}_1)$ since $\bar{\mathbf{F}}_{\bar{y}}(\mathbf{X}_1)$ can be obtained formally from $\hat{\mathbf{F}}_{\bar{y}}(\mathbf{X}_1)$ by setting $d = 1$.

Lemma I is often of assistance in evaluating the determinants. For example, if the normal rank of \mathbf{X} is unity, as is the case when \mathfrak{F} contains a single equation, \mathbf{X} can be written as $\mathbf{X}_1\mathbf{X}_2$ where \mathbf{X}_1 and \mathbf{X}_2 respectively are $q \times 1$ and $1 \times p$ matrices. In such instances the determinants of the loop-difference matrix, null loop-difference matrix, and complementary loop-difference matrix can be expressed respectively as simply $1 - \mathbf{X}_2\mathbf{C}\mathbf{X}_1$, $1 - \mathbf{X}_2\hat{\mathbf{C}}\mathbf{X}_1$, and $1 - \mathbf{X}_2\bar{\mathbf{C}}\mathbf{X}_1$. Similar simplifications can of course be exploited if \mathbf{C} , $\hat{\mathbf{C}}$, or $\bar{\mathbf{C}}$ is of unit normal rank.

With the result stated in Lemma I as motivation, we shall throughout the remainder of the paper denote $\det \mathbf{F}_{\mathfrak{F}}(\mathbf{X}_1)$, $\det \hat{\mathbf{F}}_{\mathfrak{F}}(\mathbf{X}_1)$, and $\det \bar{\mathbf{F}}_{\mathfrak{F}}(\mathbf{X}_1)$ respectively by $\det \mathbf{F}_{\mathfrak{F}}$, $\det \hat{\mathbf{F}}_{\mathfrak{F}}$, and $\det \bar{\mathbf{F}}_{\mathfrak{F}}$.

Recall that the components of Ψ' and Φ respectively are the controlled variables and controlling variables arranged in any definite orders. Thus the matrices \mathbf{X} , $\mathbf{F}_{\mathfrak{F}}(\mathbf{X}_1)$, and $\hat{\mathbf{F}}_{\mathfrak{F}}(\mathbf{X}_1)$ are not uniquely defined by the details of the two-port structure and the set of controlled-source constraints \mathfrak{F} . Nevertheless,

Lemma II: $\det \mathbf{F}_{\mathfrak{F}}$ and $\det \hat{\mathbf{F}}_{\mathfrak{F}}$ are invariant with respect to the ordering of the components of Ψ' and Φ .

In other words, the determinants are uniquely defined by the details of the two-port and \mathfrak{F} . The proof of this important result is straightforward: Let Φ' and Ψ' respectively denote vectors obtained from Φ and Ψ' by any reordering of the components. Then $\Phi' = \mathbf{U}\Phi$ and $\Psi' = \mathbf{V}\Psi'$, where \mathbf{U} and \mathbf{V} are nonsingular matrices. In terms of Φ' and Ψ' the controlled-source constraints read $\Psi' = \mathbf{X}'\Phi'$ where $\mathbf{X}' = \mathbf{V}\mathbf{X}\mathbf{U}^{-1}$, and the equations corresponding to (3) and (4) are

$$y_2 = dy_1 + \mathbf{B}'\Psi' = dy_1 + \mathbf{B}\mathbf{V}^{-1}\Psi' \quad (7)$$

$$\Phi' = \mathbf{A}'y_1 + \mathbf{C}'\Psi' = \mathbf{U}\mathbf{A}y_1 + \mathbf{U}\mathbf{C}\mathbf{V}^{-1}\Psi'. \quad (8)$$

Finally, note that for all $d \neq 0$

$$\begin{aligned} \det [\mathbf{I}_q - \mathbf{X}'(\mathbf{C}' - d^{-1}\mathbf{A}'\mathbf{B}')] &= \det [\mathbf{I}_q - \mathbf{V}\mathbf{X}\mathbf{U}^{-1}(\mathbf{U}\mathbf{C}\mathbf{V}^{-1} - d^{-1}\mathbf{U}\mathbf{A}\mathbf{B}\mathbf{V}^{-1})] \\ &= \det [\mathbf{I}_q - \mathbf{X}(\mathbf{C} - d^{-1}\mathbf{A}\mathbf{B})]. \end{aligned}$$

VI. GENERALIZATION OF BLACKMAN'S EQUATION

At this point we are in a position to state and prove the following generalization of Blackman's classical result.

Theorem I: If $w(\mathbf{0}) \neq 0$,

$$w(\mathbf{X}) = w(\mathbf{0}) \frac{\det \hat{\mathbf{F}}_{\mathfrak{F}}}{\det \mathbf{F}_{\mathfrak{F}}}.$$

6.1 Proof

Consider equation (6), which for convenience is repeated below:

$$w(\mathbf{X}) = y_2(y_1)^{-1} = w(\mathbf{0}) + \mathbf{B}[\mathbf{1}_q - \mathbf{XC}]^{-1}\mathbf{XA}. \tag{9}$$

Recall that $\text{adj}(\mathbf{M})$, the adjoint of an arbitrary square matrix \mathbf{M} of order q , is defined by

$$\mathbf{M} \text{adj}(\mathbf{M}) = \text{adj}(\mathbf{M})\mathbf{M} = \mathbf{1}_q \det \mathbf{M}.$$

Thus $w(\mathbf{X})$ can be expressed as

$$\begin{aligned} w(\mathbf{X}) &= w(\mathbf{0}) + \mathbf{B} \text{adj}(\mathbf{1}_q - \mathbf{XC})\mathbf{XA}[\det(\mathbf{1}_q - \mathbf{XC})]^{-1} \\ &= w(\mathbf{0}) \left[\frac{\det[\mathbf{1}_q - \mathbf{XC}] + w(\mathbf{0})^{-1}\mathbf{B} \text{adj}(\mathbf{1}_q - \mathbf{XC})\mathbf{XA}}{\det[\mathbf{1}_q - \mathbf{XC}]} \right]. \end{aligned}$$

Since $\det \mathbf{F}_{\bar{v}} = \det[\mathbf{1}_q - \mathbf{XC}]$, we must prove that $\det[\mathbf{1}_q - \mathbf{XC}] + w(\mathbf{0})^{-1}\mathbf{B} \text{adj}(\mathbf{1}_q - \mathbf{XC})\mathbf{XA} = \det \hat{\mathbf{F}}_{\bar{v}}$ or, more explicitly, that

$$\begin{aligned} \det[\mathbf{1}_q - \mathbf{XC}] + w(\mathbf{0})^{-1}\mathbf{B} \text{adj}(\mathbf{1}_q - \mathbf{XC})\mathbf{XA} \\ = \det[\mathbf{1}_q - \mathbf{X}(\mathbf{C} - w(\mathbf{0})^{-1}\mathbf{AB})]. \end{aligned} \tag{10}$$

We first prove the following result which will be used also in a later section.

Lemma III: Let \mathbf{G} and \mathbf{H} respectively be nonsingular matrices of orders n and m , and let \mathbf{I} and \mathbf{J} respectively denote matrices of orders $n \times m$ and $m \times n$. Then,

$$\det[\mathbf{G} + \mathbf{IH}^{-1}\mathbf{J}]\det \mathbf{H} = \det[\mathbf{H} + \mathbf{JG}^{-1}\mathbf{I}]\det \mathbf{G}.$$

The proof follows from the lemma established in Appendix A which states that $\det[\mathbf{1}_n + \mathbf{DE}] = \det[\mathbf{1}_m + \mathbf{ED}]$ for arbitrary matrices \mathbf{D} and \mathbf{E} respectively of orders $n \times m$ and $m \times n$. Taking $\mathbf{D} = \mathbf{G}^{-1}\mathbf{I}$ and $\mathbf{E} = \mathbf{H}^{-1}\mathbf{J}$, we have: $\det[\mathbf{1}_n + \mathbf{G}^{-1}\mathbf{IH}^{-1}\mathbf{J}] \det \mathbf{H} \det \mathbf{G} = \det[\mathbf{1}_m + \mathbf{H}^{-1}\mathbf{JG}^{-1}\mathbf{I}] \det \mathbf{G} \det \mathbf{H}$. Moreover, $\det[\mathbf{1}_n + \mathbf{G}^{-1}\mathbf{IH}^{-1}\mathbf{J}] \det \mathbf{G} = \det[\mathbf{G} + \mathbf{IH}^{-1}\mathbf{J}]$, and $\det[\mathbf{1}_m + \mathbf{H}^{-1}\mathbf{JG}^{-1}\mathbf{I}] \det \mathbf{H} = \det[\mathbf{H} + \mathbf{JG}^{-1}\mathbf{I}]$.

The identity (10) is a direct consequence of the following corollary of Lemma III.

Lemma IV: Let \mathbf{G} , \mathbf{I} , and \mathbf{J} respectively denote matrices of orders $n \times n$, $n \times 1$, and $1 \times n$. Then,

$$\det[\mathbf{G} + \mathbf{IJ}] = \det \mathbf{G} + \mathbf{J} \text{adj}(\mathbf{G})\mathbf{I}.$$

To establish this result, consider Lemma III with $\mathbf{H} = \mathbf{1}$ and $m = 1$. Clearly when \mathbf{G} is nonsingular, $\det[\mathbf{G} + \mathbf{IJ}] = (1 + \mathbf{JG}^{-1}\mathbf{I}) \det \mathbf{G} =$

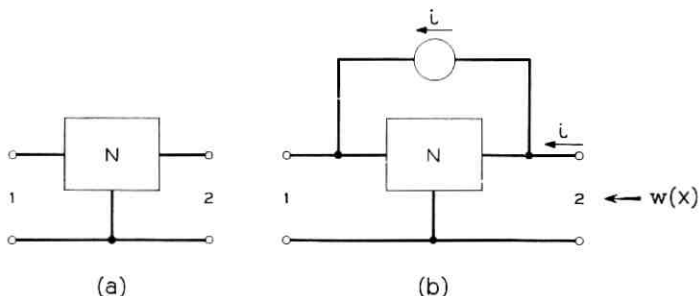


Fig. 5 — (a) Network with the open-circuit transfer impedance from port 1 to port 2 equal to $w(\mathbf{X})$; (b) modified network with driving-point impedance $w(\mathbf{X})$.

$\det \mathbf{G} + \mathbf{J} \text{adj}(\mathbf{G})\mathbf{I}$. A continuity argument of the type used in Appendix A shows that Lemma IV is valid also when \mathbf{G} is singular.

Expression (10) can be obtained from Lemma IV by taking $\mathbf{G} = (\mathbf{1}_q - \mathbf{X}\mathbf{C})$, $\mathbf{I} = \mathbf{X}\mathbf{A}$, and $\mathbf{J} = w(\mathbf{0})^{-1}\mathbf{B}$. This completes the proof of Theorem I.

6.2 Remarks Relating to Theorem I

When $w(\mathbf{X})$ is a driving-point impedance function, the null loop-difference matrix is evaluated under the condition that the current at the terminal pair at which the impedance is defined is adjusted so that the voltage at the terminal pair is zero. Thus, in this case, the null loop-difference matrix is equal to the loop-difference matrix evaluated from the network when the terminal pair is shorted. Similarly when $w(\mathbf{X})$ is a driving-point admittance function, the null loop-difference matrix is equal to the loop-difference matrix evaluated from the network when the terminal pair is open-circuited.

The conceptual and computational simplifications mentioned in the last paragraph are not valid if $w(\mathbf{X})$ is a transfer function.[†] However, by introducing an additional controlled source, the evaluation of the open-circuit transfer impedance or short-circuit transfer admittance of any three-terminal network can be reduced to the evaluation of a driving-point function. This reduction is illustrated in Fig. 5 for the transfer impedance case.

[†] Of course we assume here that $w(\mathbf{X})$ is not both a transfer function and a driving-point function associated with the two-port in Fig. 1.

6.3 *Relation of Theorem I to Earlier Work*

The expression given in Theorem I reduces to Blackman's classical result³ when it is applied to the driving-point immittance case in which \mathfrak{F} characterizes a single unilateral amplifier, and the equation is expressed in terms of physically-defined open-circuit and short-circuit return differences.

A particularly succinct derivation of an equation of this form (for the single amplifier case) was subsequently presented by Bode⁴ who exploited the properties of the determinant of the network immittance matrix. In addition Bode introduced the useful concept of return difference with respect to a two-terminal element. Signal-flow graph interpretations of Blackman's expression were later considered by Mason⁵ and by Truxal.⁶ Our proof of Theorem I is a generalization of Truxal's work.

Exploiting a suggestion by Bode,⁴ Mulligan⁷ has stated an entirely different generalization[†] of Blackman's equation for cases in which a multiplicity of nonreciprocal elements are to be considered. His result is an explicit expression obtained by repeated application of Blackman's original result.

6.4 *Generalization of Bode's Relation between Feedback and Impedance*

Theorem II: Let $\det \mathbf{F}'_{\mathfrak{F}}$ denote the determinant of the loop-difference matrix for the set of controlled-source constraints \mathfrak{F} , evaluated under the condition that $y_1 = \eta^{-1}y_2$. Let $w(\mathbf{X}) - w(\mathbf{0})$ not vanish identically in \mathbf{X} . Then $\det \mathbf{F}'_{\mathfrak{F}}$ is a linear-fractional function of η that vanishes in η if and only if $\eta = w(\mathbf{X})$.

The proof is based on Lemma IV: From the signal-flow graph obtained from Fig. 2 by adding a branch from y_2 to y_1 with transmission η^{-1} , it is clear that

$$\det \mathbf{F}'_{\mathfrak{F}} = \det \{ \mathbf{1}_q - \mathbf{X}[\mathbf{C} + \mathbf{A}\mathbf{B}(\eta - d)^{-1}] \}.$$

Using Lemma IV,

$$\det \mathbf{F}'_{\mathfrak{F}} = \det [\mathbf{1}_q - \mathbf{X}\mathbf{C}] - \mathbf{B} \operatorname{adj} (\mathbf{1}_q - \mathbf{X}\mathbf{C})\mathbf{X}\mathbf{A}(\eta - d)^{-1}$$

Thus the equation $\det \mathbf{F}'_{\mathfrak{F}} = 0$ implies that $\eta = w(\mathbf{X})$, assuming that

[†] Subsequent to preparing this paper and soliciting comments from colleagues, it came to the writer's attention that a result similar to Theorem I is contained in the recently published book by Yutze Chow and Etienne Cassagnol: *Linear Signal-Flow Graphs and Applications*, John Wiley and Sons, New York, 1962. They consider the situation that corresponds here to the special case in which \mathfrak{F} is a set of equations of the form $f_k = \gamma_k f_{k+q}$ ($k = 1, 2, \dots, q$). Their proof is considerably different from ours.

$\mathbf{B} \operatorname{adj} (\mathbf{1}_q - \mathbf{XC})\mathbf{XA} = w(\mathbf{X}) - w(\mathbf{0})$ does not vanish identically in \mathbf{X} .

Of course when $w(\mathbf{X})$ is a driving-point impedance function, the condition $y_1 = \eta^{-1}y_2$ corresponds to adding an impedance $-\eta$ in series with the one-port.⁴

6.5 The Expression for $w(\mathbf{X})$ when $w(\mathbf{0}) = 0$

Theorem I does not apply when $w(\mathbf{0}) = 0$. In fact, in all such instances the null loop-difference matrix does not exist. The following corollary which involves the complementary loop-difference matrix is of assistance in these cases.

Corollary I: If $w(\mathbf{0}) = 0$,

$$w(\mathbf{X}) = \frac{\det \bar{\mathbf{F}}_{\mathfrak{F}} - \det \mathbf{F}_{\mathfrak{F}}}{\det \mathbf{F}_{\mathfrak{F}}}.$$

This result is readily established: Here $d = 0$ in the flow graph of Fig. 2. Suppose that the branch d is replaced with one having unit transmission. Then the ratio of y_2 to y_1 for the resulting graph is $w(\mathbf{X}) + 1$. However, using Theorem I and the definition of $\bar{\mathbf{F}}_{\mathfrak{F}}$, $w(\mathbf{X}) + 1 = \det \bar{\mathbf{F}}_{\mathfrak{F}}[\det \mathbf{F}_{\mathfrak{F}}]^{-1}$.

VII. THE MATRICES $\mathbf{F}_{\mathfrak{F}\mathfrak{F}_0}$ AND $\hat{\mathbf{F}}_{\mathfrak{F}\mathfrak{F}_0}$, AND THE NOTION OF A "RESIDUAL SET OF EQUATIONS" OBTAINED FROM \mathfrak{F}

In this section we generalize the definitions of the loop-difference matrix and the null loop-difference matrix and then introduce the notion of a "residual set of equations" obtained from \mathfrak{F} , in order to both facilitate and provide a firm general basis for the subsequent derivation of some fundamental properties of $\det \mathbf{F}_{\mathfrak{F}}$ and $\det \hat{\mathbf{F}}_{\mathfrak{F}}$. The complementary loop-difference matrix need not be considered separately since it is merely a special case of the null loop-difference matrix. The material presented in this and the next section is in many respects a generalization of Bode's classical theory relating to return difference with respect to a single element.⁴

7.1 The Matrices $\mathbf{F}_{\mathfrak{F}\mathfrak{F}_0}$ and $\hat{\mathbf{F}}_{\mathfrak{F}\mathfrak{F}_0}$

The relation between y_1 and the quantities Φ , Ψ , and y_2 clearly remains unchanged if the signal-flow graph shown in Fig. 2 is replaced with the graph in Fig. 6, where \mathbf{X}_0 is an arbitrary $q \times p$ matrix. Let the loop-difference matrix and null loop-difference matrix for the branch $(\mathbf{X} - \mathbf{X}_0)$ in Fig. 6 be defined in the same manner as for \mathbf{X} in Fig. 2,

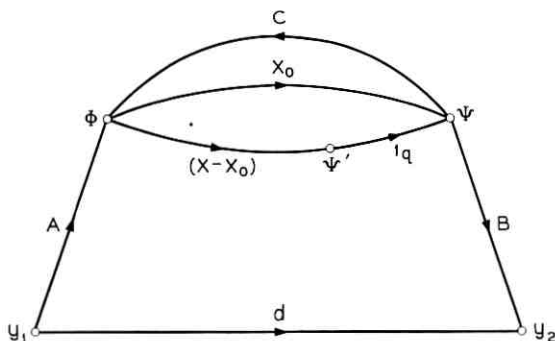


Fig. 6 — Signal-flow graph for the definition of $F_{\tilde{\mathfrak{F}}_0}$ and $\hat{F}_{\tilde{\mathfrak{F}}_0}$.

and let $\tilde{\mathfrak{F}}_0$ denote the “reference set” of equations obtained from (2) when the controlled-source matrix \mathbf{X} is set equal to \mathbf{X}_0 . We introduce

Definition I: The matrices $F_{\tilde{\mathfrak{F}}_0}$ and $\hat{F}_{\tilde{\mathfrak{F}}_0}$ respectively denote the loop-difference matrix and the null loop-difference matrix for the branch $(\mathbf{X} - \mathbf{X}_0)$ in Fig. 6. When $\mathbf{X}_0 = \mathbf{0}$, we write: $F_{\tilde{\mathfrak{F}}_0} = F_{\tilde{\mathfrak{F}}}$, $\hat{F}_{\tilde{\mathfrak{F}}_0} = \hat{F}_{\tilde{\mathfrak{F}}}$. The determinants $\det F_{\tilde{\mathfrak{F}}_0}$ and $\det \hat{F}_{\tilde{\mathfrak{F}}_0}$ respectively are referred to as the determinant of the loop-difference matrix and the determinant of the null loop-difference matrix for the set of controlled-source constraints $\tilde{\mathfrak{F}}$, with respect to the reference set $\tilde{\mathfrak{F}}_0$.

The branch $(\mathbf{X} - \mathbf{X}_0)$ in Fig. 6, which corresponds to the matrix relation $\Psi' = (\mathbf{X} - \mathbf{X}_0)\Phi$, may be interpreted as characterizing a set of controlled sources which focuses attention on the departure of the elements of \mathbf{X} from those of \mathbf{X}_0 or, equivalently, on the departure of $\tilde{\mathfrak{F}}$ from $\tilde{\mathfrak{F}}_0$.

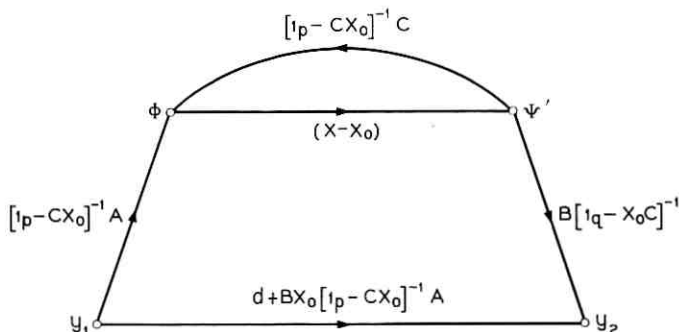
We wish to prove

Theorem III: Let $\det F_{\tilde{\mathfrak{F}}} |_{\tilde{\mathfrak{F}}=\tilde{\mathfrak{F}}_0}$ and $\det \hat{F}_{\tilde{\mathfrak{F}}} |_{\tilde{\mathfrak{F}}=\tilde{\mathfrak{F}}_0}$ respectively denote the determinants obtained from $\det F_{\tilde{\mathfrak{F}}}$ and $\det \hat{F}_{\tilde{\mathfrak{F}}}$ by replacing each element of \mathbf{X} with the corresponding element of \mathbf{X}_0 . Then,

$$\det F_{\tilde{\mathfrak{F}}_0} = \frac{\det F_{\tilde{\mathfrak{F}}}}{\det F_{\tilde{\mathfrak{F}}} |_{\tilde{\mathfrak{F}}=\tilde{\mathfrak{F}}_0}}$$

$$\det \hat{F}_{\tilde{\mathfrak{F}}_0} = \frac{\det \hat{F}_{\tilde{\mathfrak{F}}}}{\det \hat{F}_{\tilde{\mathfrak{F}}} |_{\tilde{\mathfrak{F}}=\tilde{\mathfrak{F}}_0}}.$$

Node-group Ψ' in Fig. 6 can be eliminated to obtain the flow graph in Fig. 7. For example, the branch transmission from Ψ' to Φ in Fig. 7

Fig. 7 — Flow-graph obtained from Fig. 6 by eliminating node-group Ψ .

is the transmission of the subgraph in Fig. 8. From Fig. 7 it is clear that

$$\det \mathbf{F}_{\bar{y}\bar{y}_0} = \det [\mathbf{1}_p - (\mathbf{1}_p - \mathbf{C}\mathbf{X}_0)^{-1} \mathbf{C}(\mathbf{X} - \mathbf{X}_0)]. \quad (11)$$

Thus,

$$\begin{aligned} \det \mathbf{F}_{\bar{y}\bar{y}_0} &= \det [(\mathbf{1}_p - \mathbf{C}\mathbf{X}_0)^{-1}] \det [\mathbf{1}_p - \mathbf{C}\mathbf{X}] \\ &= \frac{\det \mathbf{F}_{\bar{y}}}{\det \mathbf{F}_{\bar{y}}|_{\bar{y}=\bar{y}_0}}. \end{aligned} \quad (12)$$

Now consider $\det \hat{\mathbf{F}}_{\bar{y}\bar{y}_0}$. Since

$$w(\mathbf{X}) = w(\mathbf{0}) \frac{\det \hat{\mathbf{F}}_{\bar{y}}}{\det \mathbf{F}_{\bar{y}}} = w(\mathbf{X}_0) \frac{\det \hat{\mathbf{F}}_{\bar{y}\bar{y}_0}}{\det \mathbf{F}_{\bar{y}\bar{y}_0}} \quad (13)$$

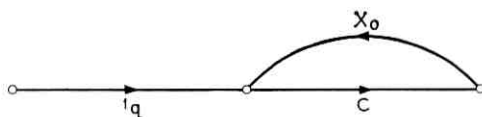
$$= [d + \mathbf{B}\mathbf{X}_0(\mathbf{1}_p - \mathbf{C}\mathbf{X}_0)^{-1} \mathbf{A}] \frac{\det \hat{\mathbf{F}}_{\bar{y}\bar{y}_0} \det \mathbf{F}_{\bar{y}}|_{\bar{y}=\bar{y}_0}}{\det \mathbf{F}_{\bar{y}}} \quad (14)$$

where (12) has been used to obtain the last expression in (14), we have

$$\det \hat{\mathbf{F}}_{\bar{y}} = [1 + d^{-1} \mathbf{B}\mathbf{X}_0(\mathbf{1}_p - \mathbf{C}\mathbf{X}_0)^{-1} \mathbf{A}] \det \hat{\mathbf{F}}_{\bar{y}\bar{y}_0} \det \mathbf{F}_{\bar{y}}|_{\bar{y}=\bar{y}_0} \quad (15)$$

$$= \{\det [\mathbf{1}_p - \mathbf{C}\mathbf{X}_0] + d^{-1} \mathbf{B}\mathbf{X}_0 \text{adj} (\mathbf{1}_p - \mathbf{C}\mathbf{X}_0) \mathbf{A}\} \det \hat{\mathbf{F}}_{\bar{y}\bar{y}_0}. \quad (16)$$

Although it is assumed in (13) that $\det \mathbf{F}_{\bar{y}}$ and $\det \mathbf{F}_{\bar{y}\bar{y}_0}$ do not vanish

Fig. 8 — Subgraph for evaluation of transmission from Ψ' to Φ in Fig. 7.

identically in the complex-frequency variable, it can readily be shown with a continuity argument that the validity of (16) does not depend upon these assumptions. According to Lemma IV, the expression within the branches in (16) is equal to $\det [\mathbf{I}_p - \hat{\mathbf{C}}\mathbf{X}_0]$. Therefore

$$\det \hat{\mathbf{F}}_{\mathfrak{F}\mathfrak{F}_0} = \frac{\det \hat{\mathbf{F}}_{\mathfrak{F}}}{\det \hat{\mathbf{F}}_{\mathfrak{F}}|_{\mathfrak{F}=\mathfrak{F}_0}}$$

7.2 The "residual set of equations"

Recall that the set \mathfrak{F} is the collection of q equations:

$$f_k = \sum_{i=q+1}^{q+p} \gamma_{ki} f_i \quad (k = 1, 2, \dots, q). \quad (17)$$

Let \mathcal{R} denote an arbitrary subset of \mathcal{R}_0 , the set of all ordered pairs of integers (j, i) such that $1 \leq j \leq q, q+1 \leq i \leq q+p$. Let $\sum_{\mathcal{R}}$ denote a sum over all integers i such that $(k, i) \in \mathcal{R}$ [i.e., such that (k, i) is an element of \mathcal{R}], and denote by \mathcal{R}^* the complement of \mathcal{R} with respect to \mathcal{R}_0 . Then it is certainly true that

$$f_k = f'_k + \sum_{\mathcal{R}^*} \gamma_{ki} f_i \quad (k = 1, 2, \dots, q)$$

where

$$f'_k = \sum_{\mathcal{R}} \gamma_{ki} f_i \quad (k = 1, 2, \dots, q). \quad (18)$$

In accordance with the controlled-source interpretation of (17), the f'_k are the contributions to the controlled-source outputs associated with the subset of coefficients $\{\gamma_{ki} | (k, i) \in \mathcal{R}\}$. Let $\mathfrak{F} \cdot \mathcal{R}$ denote the set of r equations ($r \leq q$) obtained from (18) by omitting all equations of the form $f'_k = 0$. The set $\mathfrak{F} \cdot \mathcal{R}$ is referred to as a residual set of equations obtained from \mathfrak{F} .

VIII. THEOREMS CONCERNING $\det \mathbf{F}_{\mathfrak{F} \cdot \mathcal{R}}$ AND $\det \hat{\mathbf{F}}_{\mathfrak{F} \cdot \mathcal{R}}$

To each choice of \mathcal{R} there corresponds a signal-flow graph characterization of the relation between y_1 and y_2 of the type shown in Fig. 2 where the elements of the flow matrices analogous to \mathbf{A} , \mathbf{B} , \mathbf{C} , and d are independent of those literal coefficients γ_{ki} for which $(k, i) \in \mathcal{R}$, and where the branch analogous to \mathbf{X} is associated with the set of equations $\mathfrak{F} \cdot \mathcal{R}$. Accordingly, each choice of \mathcal{R} defines a set of controlled-source interpretable equations $\mathfrak{F} \cdot \mathcal{R}$, a pair of determinants $\det \mathbf{F}_{\mathfrak{F} \cdot \mathcal{R}}$ and $\det \hat{\mathbf{F}}_{\mathfrak{F} \cdot \mathcal{R}}$, and an initial state of the physical system [i.e., a system obtained

by setting $\gamma_{ki} = 0$ for all (k,i) contained in \mathcal{R} . The primary purpose of this section is to derive some fundamental results that relate $\det \mathbf{F}_{\mathfrak{F} \cdot \mathcal{R}_1}$ and $\det \hat{\mathbf{F}}_{\mathfrak{F} \cdot \mathcal{R}_1}$ respectively to $\det \mathbf{F}_{\mathfrak{F} \cdot \mathcal{R}_2}$ and $\det \hat{\mathbf{F}}_{\mathfrak{F} \cdot \mathcal{R}_2}$, where \mathcal{R}_1 and \mathcal{R}_2 are any two subsets of \mathcal{R}_0 . The set $\mathfrak{F} \cdot \mathcal{R}_0 = \mathfrak{F}$ is of interest here only in that $\mathfrak{F} \cdot \mathcal{R}_1$ and $\mathfrak{F} \cdot \mathcal{R}_2$ are obtained from it in a prescribed manner.

Our first objective is to relate $\det \mathbf{F}_{\mathfrak{F} \cdot \mathcal{R}}$ and $\det \hat{\mathbf{F}}_{\mathfrak{F} \cdot \mathcal{R}}$ respectively to $\det \mathbf{F}_{\mathfrak{F}}$ and $\det \hat{\mathbf{F}}_{\mathfrak{F}}$. Consider

Lemma V: Let \mathbf{A}' , \mathbf{B}' , \mathbf{C}' , and d' respectively denote the flow matrices in Fig. 7 which correspond to \mathbf{A} , \mathbf{B} , \mathbf{C} , and d in Fig. 2. Let Φ , Ψ' , \mathbf{A}' , \mathbf{B}' , and \mathbf{C}' be partitioned as follows:

$$\Phi = \begin{bmatrix} \Phi_1 \\ \Phi_2 \end{bmatrix} \begin{matrix} s \\ (p-s) \end{matrix}, \quad \Psi' = \begin{bmatrix} \Psi'_1 \\ \Psi'_2 \end{bmatrix} \begin{matrix} r \\ (q-r) \end{matrix}, \quad \mathbf{A}' = \begin{bmatrix} \mathbf{A}'_1 \\ \mathbf{A}'_2 \end{bmatrix} \begin{matrix} s \\ (p-s) \end{matrix}$$

$$\mathbf{B}' = \begin{bmatrix} \mathbf{B}'_1 & \mathbf{B}'_2 \end{bmatrix} \begin{matrix} r \\ (q-r) \end{matrix}, \quad \mathbf{C}' = \begin{bmatrix} \mathbf{C}'_{11} & \mathbf{C}'_{12} \\ \mathbf{C}'_{21} & \mathbf{C}'_{22} \end{bmatrix} \begin{matrix} s \\ (p-s) \end{matrix}$$

$$r \quad (q-r)$$

and let all of the nonzero elements of $(\mathbf{X} - \mathbf{X}_0)$ be restricted to \mathbf{X}_{rs} , the $r \times s$ submatrix standing in the upper left-hand corner. Then $\det \mathbf{F}_{\mathfrak{F} \cdot \mathcal{R}_0}$ and $\det \hat{\mathbf{F}}_{\mathfrak{F} \cdot \mathcal{R}_0}$ respectively are equal to the determinant of the loop-difference matrix and the determinant of the null loop-difference matrix for the branch \mathbf{X}_{rs} in the flow graph of Fig. 9.

The proof follows at once from the expressions for $\det \mathbf{F}_{\mathfrak{F} \cdot \mathcal{R}_0}$ and $\det \hat{\mathbf{F}}_{\mathfrak{F} \cdot \mathcal{R}_0}$ in terms of d' and the submatrices of \mathbf{A}' , \mathbf{B}' and \mathbf{C}' . The details are omitted.

Suppose now that the elements of Ψ' and Φ are chosen so that all of the coefficients $\gamma_{ki} [(k,i) \in \mathcal{R}]$ are contained in the $r \times s$ submatrix in the

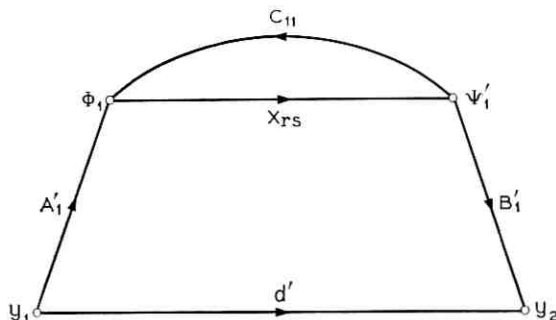


Fig. 9 — Flow-graph relevant to Lemma V.

upper left-hand corner of \mathbf{X} , where s is the smallest integer for which this is possible, and that \mathbf{X}_0 in Fig. 7 is obtained from \mathbf{X} by replacing the $\gamma_{ki}[(k,i)\varepsilon\mathcal{R}]$ by zeros. It then follows from Lemmas II and V that $\det \mathbf{F}_{\mathfrak{F}\mathcal{R}} = \det \mathbf{F}_{\mathfrak{F}\mathcal{R}_0}$ and $\det \hat{\mathbf{F}}_{\mathfrak{F}\mathcal{R}} = \det \hat{\mathbf{F}}_{\mathfrak{F}\mathcal{R}_0}$. Thus a direct application of Theorem III at this point proves

Theorem IV:

$$\det \mathbf{F}_{\mathfrak{F}\cdot\mathcal{R}} = \frac{\det \mathbf{F}_{\mathfrak{F}}}{\det \mathbf{F}_{\mathfrak{F}} \Big|_{\substack{\gamma_{ki} = 0 \\ (k,i)\varepsilon\mathcal{R}}}}$$

$$\det \hat{\mathbf{F}}_{\mathfrak{F}\cdot\mathcal{R}} = \frac{\det \hat{\mathbf{F}}_{\mathfrak{F}}}{\det \hat{\mathbf{F}}_{\mathfrak{F}} \Big|_{\substack{\gamma_{ki} = 0 \\ (k,i)\varepsilon\mathcal{R}}}}.$$

Note that $\det \mathbf{F}_{\mathfrak{F}\cdot\mathcal{R}}$ and $\det \hat{\mathbf{F}}_{\mathfrak{F}\cdot\mathcal{R}}$ are linear-fractional functions in each of the $\gamma_{ki}[(k,i)\varepsilon\mathcal{R}^*]$. This fact is often of assistance in evaluating the determinants.

Now consider $\det \mathbf{F}_{\mathfrak{F}} \Big|_{\substack{\gamma_{ki} = 0 \\ (k,i)\varepsilon\mathcal{R}}}$ and $\det \hat{\mathbf{F}}_{\mathfrak{F}} \Big|_{\substack{\gamma_{ki} = 0 \\ (k,i)\varepsilon\mathcal{R}}}$.

Recall that these quantities are respectively the determinant of the loop-difference matrix and the determinant of the null loop-difference matrix for the branch \mathbf{X} in Fig. 2 when $\gamma_{ki} = 0[(k,i)\varepsilon\mathcal{R}]$. Consequently it follows from Lemma II, Lemma V, and the significance of $\mathfrak{F}\cdot\mathcal{R}^*$ that:

$$\begin{aligned} \det \mathbf{F}_{\mathfrak{F}} \Big|_{\substack{\gamma_{ki} = 0 \\ (k,i)\varepsilon\mathcal{R}}} &= \det \mathbf{F}_{\mathfrak{F}\cdot\mathcal{R}^*} \Big|_{\substack{\gamma_{ki} = 0 \\ (k,i)\varepsilon\mathcal{R}}} \\ &= \det \mathbf{F}_{\mathfrak{F}\cdot\mathcal{R}^*} \Big|_{\substack{\gamma_{ki} = 0 \\ (k,i)\varepsilon\mathcal{R}}} \end{aligned} \tag{19}$$

$$\begin{aligned} \det \hat{\mathbf{F}}_{\mathfrak{F}} \Big|_{\substack{\gamma_{ki} = 0 \\ (k,i)\varepsilon\mathcal{R}}} &= \det \hat{\mathbf{F}}_{\mathfrak{F}\cdot\mathcal{R}^*} \Big|_{\substack{\gamma_{ki} = 0 \\ (k,i)\varepsilon\mathcal{R}}} \\ &= \det \hat{\mathbf{F}}_{\mathfrak{F}\cdot\mathcal{R}^*} \Big|_{\substack{\gamma_{ki} = 0 \\ (k,i)\varepsilon\mathcal{R}}} \end{aligned} \tag{20}$$

where $\det \mathbf{F}_{\mathfrak{F}\cdot\mathcal{R}^*} \Big|_{\substack{\gamma_{ki} = 0 \\ (k,i)\varepsilon\mathcal{R}}}$ and $\det \hat{\mathbf{F}}_{\mathfrak{F}\cdot\mathcal{R}^*} \Big|_{\substack{\gamma_{ki} = 0 \\ (k,i)\varepsilon\mathcal{R}}}$ respectively

are equal to $\det \mathbf{F}_{\mathfrak{F}\cdot\mathcal{R}^*}$ and $\det \hat{\mathbf{F}}_{\mathfrak{F}\cdot\mathcal{R}^*}$ evaluated from the flow graph or directly from the circuit under the condition that $\gamma_{ki} = 0$ for all (k,i) contained in \mathcal{R} . Theorem IV and identities (19) and (20) imply

Theorem V: Let \mathcal{R}_1 and \mathcal{R}_2 denote two arbitrary subsets of \mathcal{R}_0 . Then

$$\frac{\det \mathbf{F}_{\mathfrak{F} \cdot \mathcal{R}_1}}{\det \mathbf{F}_{\mathfrak{F} \cdot \mathcal{R}_2}} = \frac{\det \mathbf{F}_{\mathfrak{F} \cdot \mathcal{R}_2^*} \left| \begin{array}{l} \gamma_{ki} = 0 \\ (k,i) \in \mathcal{R}_2 \end{array} \right.}{\det \mathbf{F}_{\mathfrak{F} \cdot \mathcal{R}_1^*} \left| \begin{array}{l} \gamma_{ki} = 0 \\ (k,i) \in \mathcal{R}_1 \end{array} \right.}$$

$$\frac{\det \hat{\mathbf{F}}_{\mathfrak{F} \cdot \mathcal{R}_1}}{\det \hat{\mathbf{F}}_{\mathfrak{F} \cdot \mathcal{R}_2}} = \frac{\det \hat{\mathbf{F}}_{\mathfrak{F} \cdot \mathcal{R}_2^*} \left| \begin{array}{l} \gamma_{ki} = 0 \\ (k,i) \in \mathcal{R}_2 \end{array} \right.}{\det \hat{\mathbf{F}}_{\mathfrak{F} \cdot \mathcal{R}_1^*} \left| \begin{array}{l} \gamma_{ki} = 0 \\ (k,i) \in \mathcal{R}_1 \end{array} \right.}.$$

We wish now to focus attention on the particular situation in which \mathcal{R}_1 and \mathcal{R}_2 are disjoint. We shall prove

Theorem VI: Let \mathcal{R}_1 and \mathcal{R}_2 be any two disjoint subsets of \mathcal{R}_0 . Then

$$\frac{\det \mathbf{F}_{\mathfrak{F} \cdot \mathcal{R}_1}}{\det \mathbf{F}_{\mathfrak{F} \cdot \mathcal{R}_2}} = \frac{\det \mathbf{F}_{\mathfrak{F} \cdot \mathcal{R}_1} \left| \begin{array}{l} \gamma_{ki} = 0 \\ (k,i) \in \mathcal{R}_2 \end{array} \right.}{\det \mathbf{F}_{\mathfrak{F} \cdot \mathcal{R}_2} \left| \begin{array}{l} \gamma_{ki} = 0 \\ (k,i) \in \mathcal{R}_1 \end{array} \right.}$$

$$\frac{\det \hat{\mathbf{F}}_{\mathfrak{F} \cdot \mathcal{R}_1}}{\det \hat{\mathbf{F}}_{\mathfrak{F} \cdot \mathcal{R}_2}} = \frac{\det \hat{\mathbf{F}}_{\mathfrak{F} \cdot \mathcal{R}_1} \left| \begin{array}{l} \gamma_{ki} = 0 \\ (k,i) \in \mathcal{R}_2 \end{array} \right.}{\det \hat{\mathbf{F}}_{\mathfrak{F} \cdot \mathcal{R}_2} \left| \begin{array}{l} \gamma_{ki} = 0 \\ (k,i) \in \mathcal{R}_1 \end{array} \right.}.$$

Consider Theorem V. Observe that here $\mathfrak{F} \cdot \mathcal{R}_1$ and $\mathfrak{F} \cdot \mathcal{R}_2$ respectively can be regarded as residual sets of equations obtained from $\mathfrak{F} \cdot \mathcal{R}_2^*$ and $\mathfrak{F} \cdot \mathcal{R}_1^*$. Formally, in accordance with the notation introduced earlier, $\mathfrak{F} \cdot \mathcal{R}_1 = (\mathfrak{F} \cdot \mathcal{R}_2^*) \cdot \mathcal{R}_1$ and $\mathfrak{F} \cdot \mathcal{R}_2 = (\mathfrak{F} \cdot \mathcal{R}_1^*) \cdot \mathcal{R}_2$. Using Theorem IV

$$\det \mathbf{F}_{\mathfrak{F} \cdot \mathcal{R}_1} = \frac{\det \mathbf{F}_{\mathfrak{F} \cdot \mathcal{R}_2^*}}{\det \mathbf{F}_{\mathfrak{F} \cdot \mathcal{R}_2^*} \left| \begin{array}{l} \gamma_{ki} = 0 \\ (k,i) \in \mathcal{R}_1 \end{array} \right.}}.$$

Hence

$$\frac{\det \mathbf{F}_{\mathfrak{F} \cdot \mathcal{R}_2^*} \left| \begin{array}{l} \gamma_{ki} = 0 \\ (k,i) \in \mathcal{R}_2 \end{array} \right.}{\det \mathbf{F}_{\mathfrak{F} \cdot \mathcal{R}_2^*} \left| \begin{array}{l} \gamma_{ki} = 0 \\ (k,i) \in \mathcal{R}_1 \cup \mathcal{R}_2 \end{array} \right.}} = \det \mathbf{F}_{\mathfrak{F} \cdot \mathcal{R}_1} \left| \begin{array}{l} \gamma_{ki} = 0 \\ (k,i) \in \mathcal{R}_2 \end{array} \right. \quad (21)$$

$$= \det \mathbf{F}_{\mathfrak{F} \cdot \mathcal{R}_1} \left| \begin{array}{l} \gamma_{ki} = 0 \\ (k,i) \in \mathcal{R}_2 \end{array} \right. .$$

Next let $\mathcal{R}_3 = (\mathcal{R}_1 \cup \mathcal{R}_2)^*$ and note that $\mathfrak{F} \cdot \mathcal{R}_3 = (\mathfrak{F} \cdot \mathcal{R}_1^*) \cdot \mathcal{R}_3 = (\mathfrak{F} \cdot \mathcal{R}_2^*) \cdot \mathcal{R}_3$. Again using Theorem IV

$$\det \mathbf{F}_{\mathfrak{F} \cdot \mathcal{R}_3} = \frac{\det \mathbf{F}_{\mathfrak{F} \cdot \mathcal{R}_1^*}}{\det \mathbf{F}_{\mathfrak{F} \cdot \mathcal{R}_1^*} \left| \begin{array}{l} \gamma_{ki} = 0 \\ (k,i) \in \mathcal{R}_3 \end{array} \right.} = \frac{\det \mathbf{F}_{\mathfrak{F} \cdot \mathcal{R}_2^*}}{\det \mathbf{F}_{\mathfrak{F} \cdot \mathcal{R}_2^*} \left| \begin{array}{l} \gamma_{ki} = 0 \\ (k,i) \in \mathcal{R}_3 \end{array} \right.}.$$

Thus

$$\begin{aligned} \det \mathbf{F}_{\mathfrak{F} \cdot \mathcal{R}_3} \left| \begin{array}{l} \gamma_{ki} = 0 \\ (k,i) \in \mathcal{R}_1 \cup \mathcal{R}_2 \end{array} \right. &= \det \mathbf{F}_{\mathfrak{F} \cdot \mathcal{R}_1^*} \left| \begin{array}{l} \gamma_{ki} = 0 \\ (k,i) \in \mathcal{R}_1 \cup \mathcal{R}_2 \end{array} \right. \\ &= \det \mathbf{F}_{\mathfrak{F} \cdot \mathcal{R}_2^*} \left| \begin{array}{l} \gamma_{ki} = 0 \\ (k,i) \in \mathcal{R}_1 \cup \mathcal{R}_2 \end{array} \right. \end{aligned} \tag{22}$$

Finally, from (21), (22), (19) and the fact that (21) remains valid when the subscripts 1 and 2 are interchanged, it follows directly that

$$\frac{\det \mathbf{F}_{\mathfrak{F} \cdot \mathcal{R}_2^*} \left| \begin{array}{l} \gamma_{ki} = 0 \\ (k,i) \in \mathcal{R}_2 \end{array} \right.}{\det \mathbf{F}_{\mathfrak{F} \cdot \mathcal{R}_1^*} \left| \begin{array}{l} \gamma_{ki} = 0 \\ (k,i) \in \mathcal{R}_1 \end{array} \right.} = \frac{\det \mathbf{F}_{\mathfrak{F} \cdot \mathcal{R}_1} \left| \begin{array}{l} \gamma_{ki} = 0 \\ (k,i) \in \mathcal{R}_2 \end{array} \right.}{\det \mathbf{F}_{\mathfrak{F} \cdot \mathcal{R}_2} \left| \begin{array}{l} \gamma_{ki} = 0 \\ (k,i) \in \mathcal{R}_1 \end{array} \right.}.$$

when \mathcal{R}_1 and \mathcal{R}_2 are disjoint. It is obvious at this point that a similar argument suffices to establish the second identity stated in the theorem.

It is worth stating explicitly the following direct corollary of Theorem VI.

Corollary II: Let N_1 and N_2 be two arbitrary disjoint subsets of $\{1, 2, \dots, q\}$. Let \mathfrak{F}_1 and \mathfrak{F}_2 respectively denote the subsets of equations obtained from \mathfrak{F} by including only those for which $k \in N_1$ and $k \in N_2$. Then

$$\begin{aligned} \frac{\det \mathbf{F}_{\mathfrak{F}_1} \left| \begin{array}{l} \gamma_{ki} = 0 \\ k \in N_2 \end{array} \right.}{\det \mathbf{F}_{\mathfrak{F}_2} \left| \begin{array}{l} \gamma_{ki} = 0 \\ k \in N_1 \end{array} \right.} &= \frac{\det \mathbf{F}_{\mathfrak{F}_1} \left| \begin{array}{l} \gamma_{ki} = 0 \\ k \in N_2 \end{array} \right.}{\det \mathbf{F}_{\mathfrak{F}_2} \left| \begin{array}{l} \gamma_{ki} = 0 \\ k \in N_1 \end{array} \right.} \\ \frac{\det \hat{\mathbf{F}}_{\mathfrak{F}_1} \left| \begin{array}{l} \gamma_{ki} = 0 \\ k \in N_2 \end{array} \right.}{\det \hat{\mathbf{F}}_{\mathfrak{F}_2} \left| \begin{array}{l} \gamma_{ki} = 0 \\ k \in N_1 \end{array} \right.} &= \frac{\det \hat{\mathbf{F}}_{\mathfrak{F}_1} \left| \begin{array}{l} \gamma_{ki} = 0 \\ k \in N_2 \end{array} \right.}{\det \hat{\mathbf{F}}_{\mathfrak{F}_2} \left| \begin{array}{l} \gamma_{ki} = 0 \\ k \in N_1 \end{array} \right.}. \end{aligned}$$

The corollary is of utility in evaluating $\det \mathbf{F}_{\mathfrak{F}_2}$ and $\det \hat{\mathbf{F}}_{\mathfrak{F}_2}$ when $\det \mathbf{F}_{\mathfrak{F}_1}$ and $\det \hat{\mathbf{F}}_{\mathfrak{F}_1}$ are known, since

$$\det \mathbf{F}_{\mathfrak{F}_2} \left| \begin{array}{l} \gamma_{ki} = 0 \\ k \in N_1 \end{array} \right. \quad \text{and} \quad \det \hat{\mathbf{F}}_{\mathfrak{F}_2} \left| \begin{array}{l} \gamma_{ki} = 0 \\ k \in N_1 \end{array} \right.$$

are often considerably easier to evaluate than $\det \mathbf{F}_{\mathfrak{F}_2}$ and $\det \hat{\mathbf{F}}_{\mathfrak{F}_2}$. Of course similar remarks apply to Theorem VI. Frequently Theorem V

also is useful in this respect. These results are generalizations of a theorem due to Bode.⁴

The following factorization theorem can be obtained by repeated applications of Theorem IV.

Theorem VII: Let $\{\mathcal{R}_1, \mathcal{R}_2, \dots, \mathcal{R}_n\}$ denote any collection of disjoint subsets of \mathcal{R}_0 such that $\bigcup_{i=1}^n \mathcal{R}_i = \mathcal{R}_0$. Then

$$\det \mathbf{F}_{\mathfrak{F}} = \det \mathbf{F}_{\mathfrak{F}, \mathcal{R}_1} \prod_{i=2}^n \det \mathbf{F}_{\mathfrak{F}, \mathcal{R}_i} \left\| \begin{array}{l} \gamma_{ki} = 0 \\ (k, i) \in \mathcal{R}_i' \end{array} \right.$$

$$\det \hat{\mathbf{F}}_{\mathfrak{F}} = \det \hat{\mathbf{F}}_{\mathfrak{F}, \mathcal{R}_1} \prod_{i=2}^n \det \hat{\mathbf{F}}_{\mathfrak{F}, \mathcal{R}_i} \left\| \begin{array}{l} \gamma_{ki} = 0 \\ (k, i) \in \mathcal{R}_i' \end{array} \right.$$

where $\mathcal{R}_i' = \bigcup_{j=1}^{i-1} \mathcal{R}_j$.

Observe that $\det \mathbf{F}_{\mathfrak{F}}$ is expressed as a product of ordinary loop-differences^{4,5} when each $\mathcal{R}_i (i = 1, 2, \dots, n)$ contains a single element.

IX. SOME SIMPLE APPLICATIONS OF THE THEORY

The first three examples relate to a specific vacuum-tube circuit that has been considered by Truxal.⁶ He presents a detailed classical flow-graph analysis.

9.1 An Application of Theorem I

To illustrate the application of Theorem I we shall compute the driving-point impedance at port 2 of the circuit shown in Fig. 10 when port 1 is short-circuited. The pertinent linear incremental model is shown in Fig. 10. It is assumed that $\mu_1 = \mu_2 = 20$, $r_{p1} = r_{p2} = 10$, and $R_{L2} = R_{L1}R(R_{L1} + R)^{-1} = 200$.

Here we choose as the set \mathfrak{F} the two equations

$$e_a = \mu_1 e_{\varrho 1}$$

$$e_b = \mu_2 e_{\varrho 2}$$

Let $\Psi^t = [e_a, e_b]^t$ and $\Phi = [e_{\varrho 1}, e_{\varrho 2}]^t$, where the superscript t indicates transposition. Hence $\mathbf{X} = \text{diag} [\mu_1, \mu_2]$. It is a trivial matter to show that

$$w(0) = \frac{440R_k + 4200}{4.2R_k + 441}.$$

Recall that \mathbf{C} is defined here by $\Phi = \mathbf{C}\Psi$ where the elements of Ψ are treated as independent variables and port 2 is open-circuited. Similarly $\Phi = \hat{\mathbf{C}}\Psi$ when port 2 is short-circuited. An elementary analysis yields:

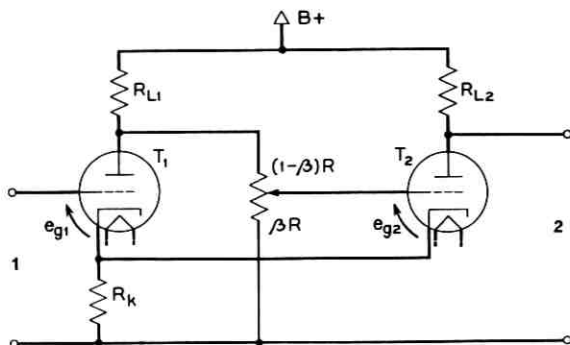


Fig. 10 — Circuit diagram for the example.

$$\mathbf{C} = \begin{bmatrix} \frac{-2.1R_k}{4.2R_k + 441} & \frac{-2.1R_k}{4.2R_k + 441} \\ \frac{-R_k(2\beta + 2.1) - 420\beta}{4.2R_k + 441} & \frac{R_k(2\beta - 2.1)}{4.2R_k + 441} \end{bmatrix} \quad (23)$$

$$\hat{\mathbf{C}} = \begin{bmatrix} \frac{-R_k}{22R_k + 210} & \frac{-21R_k}{22R_k + 210} \\ \frac{-R_k(20\beta + 1) - 200\beta}{22R_k + 210} & \frac{21R_k \left(\frac{20}{21} \beta - 1 \right)}{22R_k + 210} \end{bmatrix}.$$

The determinants $\det \mathbf{F}_{\beta} = \det [\mathbf{I}_2 - \mathbf{C}\mathbf{X}]$ and $\det \hat{\mathbf{F}}_{\beta} = \det [\mathbf{I}_2 - \hat{\mathbf{C}}\mathbf{X}]$ can be evaluated in a particularly simple manner by exploiting the fact that they are known at the outset to be linear-fractional functions in R_k . We find that

$$\det \mathbf{F}_{\beta} = \frac{(88.2 - 840\beta)R_k + 441}{4.2R_k + 441} \quad (24)$$

$$\det \hat{\mathbf{F}}_{\beta} = \frac{(46.2 - 840\beta)R_k + 210}{22R_k + 210}$$

when $\mu_1 = \mu_2 = 20$. Therefore

$$\begin{aligned}
 w(\text{diag } [20,20]) &= w(0) \frac{\det \hat{\mathbf{F}}_{\beta}}{\det \mathbf{F}_{\beta}} \\
 &= 200 \frac{(46.2 - 840\beta)R_k + 210}{(88.2 - 840\beta)R_k + 441}.
 \end{aligned}$$

9.2 An Application of Corollary I

In this section let $w(\mathbf{X})$ be the (port 1 to port 2) open-circuit voltage transfer function for the circuit in Fig. 10, and let \mathfrak{F} , Ψ , and Φ be

chosen as in the last section. Our objective here is to evaluate the transfer function in accordance with Corollary I which is obviously relevant.

Since \mathbf{C} and $\det \mathbf{F}_{\mathfrak{F}}$ have already been evaluated, consider the determination of the pertinent flow matrices \mathbf{A} and \mathbf{B} . By inspection of the circuit, $\mathbf{A} = [1, 0]^t$ and a simple calculation yields

$$\mathbf{B} = \begin{bmatrix} \frac{2R_k}{4.2R_k + 441}, & \frac{-2(R_k + 210)}{4.2R_k + 441} \end{bmatrix}.$$

It follows that

$$\begin{aligned} \det \bar{\mathbf{F}}_{\mathfrak{F}} &= \det [\mathbf{I}_2 - (\mathbf{C} - \mathbf{A}\mathbf{B})\mathbf{X}] \\ &= \det \begin{bmatrix} \frac{86.2R_k + 441}{4.2R_k + 441} & \frac{2R_k - 20(420)}{4.2R_k + 441} \\ \frac{R_k(40\beta + 42) + 8400\beta}{4.2R_k + 441} & \frac{(46.2 - 40\beta)R_k + 441}{4.2R_k + 441} \end{bmatrix} \\ &= \frac{(928.2 - 840\beta)R_k + 441 + 160,000\beta}{4.2R_k + 441}. \end{aligned} \quad (25)$$

Therefore from Corollary I, (24), and (25)

$$v(\text{diag } [20, 20]) = \frac{840R_k + 160,000\beta}{(88.2 - 840\beta)R_k + 441}.$$

9.3 An Application of Corollary II

Consider the model in Fig. 11. Here let \mathfrak{F} denote the set of equations

$$e_a = \mu_1 e_{g1}$$

$$e_b = \mu_2 e_{g2}$$

$$e_k = R_k i_k$$

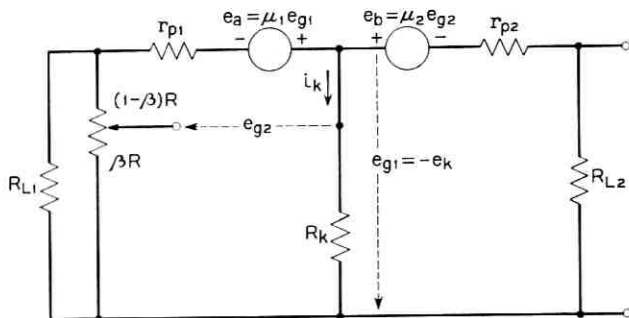


Fig. 11 — Linear model for the network in Fig. 10 when port 1 is short-circuited.

and let \mathfrak{F}_1 and \mathfrak{F}_2 respectively denote the subsets

$$\begin{aligned} e_a &= \mu_1 e_{a1} \\ & , \quad e_k = R_k i_k . \\ e_b &= \mu_2 e_{b2} \end{aligned}$$

Truxal shows⁶ that $\det \mathbf{F}_{\mathfrak{F}_2} = 1 - [-0.2 + (40/21)\beta]R_k$ when the network parameters have the values given in Section 9.1. We wish to determine the corresponding expression for $\det \mathbf{F}_{\mathfrak{F}_1}$ using Corollary II.

By inspection of Fig. 11

$$\det \mathbf{F}_{\mathfrak{F}_1} \parallel^{R_k=0} = 1, \quad \det \mathbf{F}_{\mathfrak{F}_2} \parallel^{\mu_1=\mu_2=0} = 1 + \frac{1}{105} R_k .$$

Thus

$$\det \mathbf{F}_{\mathfrak{F}_1} = \frac{(88.2 - 840\beta)R_k + 441}{4.2R_k + 441}$$

which of course is identical to the right-hand side of (24).

9.4 A Flow Graph Demonstration of Corollary II

Corollary II can be demonstrated by considering the flow graph in Fig. 12. Let the determinants of the loop-difference matrices for the branches \mathbf{Y}_1 and \mathbf{Y}_2 respectively be denoted by $\det \mathbf{F}_{\mathfrak{F}_1}$ and $\det \mathbf{F}_{\mathfrak{F}_2}$. Straightforward evaluation shows that

$$\det \mathbf{F}_1 = \det \{ \mathbf{1}_m - \mathbf{Y}_1[\mathbf{C}_1 + \mathbf{K}(\mathbf{1}_n - \mathbf{Y}_2\mathbf{C}_2)^{-1}\mathbf{Y}_2\mathbf{L}] \}$$

$$\det \mathbf{F}_{\mathfrak{F}_2} = \det \{ \mathbf{1}_n - \mathbf{Y}_2[\mathbf{C}_2 + \mathbf{L}(\mathbf{1}_m - \mathbf{Y}_1\mathbf{C}_1)^{-1}\mathbf{Y}_1\mathbf{K}] \}$$

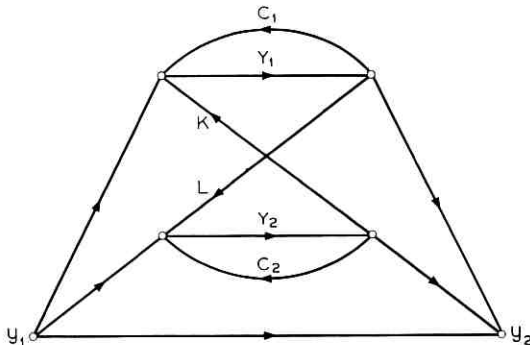


Fig. 12 — Flow-graph relating to the validity of Corollary II.

where n and m respectively are the number of rows of \mathbf{Y}_2 and the number of rows of \mathbf{Y}_1 .

The corollary implies that

$$\frac{\det \{ \mathbf{1}_m - \mathbf{Y}_1[\mathbf{C}_1 + \mathbf{K}(\mathbf{1}_n - \mathbf{Y}_2\mathbf{C}_2)^{-1}\mathbf{Y}_2\mathbf{L}] \}}{\det \{ \mathbf{1}_n - \mathbf{Y}_2[\mathbf{C}_2 + \mathbf{L}(\mathbf{1}_m - \mathbf{Y}_1\mathbf{C}_1)^{-1}\mathbf{Y}_1\mathbf{K}] \}} = \frac{\det [\mathbf{1}_m - \mathbf{Y}_1\mathbf{C}_1]}{\det [\mathbf{1}_n - \mathbf{Y}_2\mathbf{C}_2]}$$

which can readily be verified independently with the aid of Lemma III.

9.5 Foster's Results Via Theorem I

As a final illustration of the theory we shall outline briefly an alternative proof of Foster's well-known results concerning the realizability of two-element-kind one-ports.

Let an arbitrary passive RC one-port containing only resistors and q parallel-RC combinations with capacitance and resistance values respectively c_k and $(\alpha c_k)^{-1}$ ($k = 1, 2, \dots, q$) be characterized by the driving-point admittance function $y(s + \alpha)$, where α is real and positive. Let \mathfrak{F} denote the set of capacitor volt-ampere constraints: $i_k = sc_k e_k$ ($k = 1, 2, \dots, q$) and consider the evaluation of $y(s + \alpha)$ in accordance with Theorem I (some of the flow matrices in Fig. 2 may not exist if $\alpha = 0$).

From Theorem I and Lemma I, $y(s + \alpha)$ can be expressed as

$$y(s + \alpha) = w(\mathbf{0}) \frac{\det [\mathbf{1}_q - s\mathbf{C}_1\hat{\mathbf{C}}\mathbf{C}_1]}{\det [\mathbf{1}_q - s\mathbf{C}_1\mathbf{C}\mathbf{C}_1]} \quad (26)$$

where $\mathbf{C}_1 = \text{diag} [(c_1)^{\frac{1}{2}}, (c_2)^{\frac{1}{2}}, \dots, (c_q)^{\frac{1}{2}}]$. Observe that here $-\mathbf{C}$ and $-\hat{\mathbf{C}}$ are passive open-circuit resistance matrices. A moment's reflection shows that $\mathbf{A} = \mathbf{B}^t$. Hence $-\hat{\mathbf{C}} = -\mathbf{C} + \mathbf{D}$ where \mathbf{D} is a nonnegative definite matrix of unit rank. With the aid of (26) and the following theorem it becomes a simple matter to show that

$$y(s) = g_0 + sc_\infty + \sum_{k=1}^{q'} \frac{sa_k}{s + b_k}$$

where $g_0, c_\infty \geq 0$; $a_k, b_k \geq 0$ ($k = 1, 2, \dots, q'$).

Theorem VIII: Let \mathbf{P} and \mathbf{Q} denote two nonnegative definite hermitian matrices of order n ; let \mathbf{Q} have unit rank. Then

$$p_1 \leq r_1 \leq p_2 \leq r_2 \cdots \leq p_n \leq r_n$$

where $\{p_1, p_2, \dots, p_n\}$ is the set of eigenvalues of \mathbf{P} and $\{r_1, r_2, \dots, r_n\}$ is the set of eigenvalues of $(\mathbf{P} + \mathbf{Q})$. Further, if p_j ($j = 1, 2, \dots, n$) is an eigenvalue of \mathbf{P} of multiplicity m_j ($m_j \geq 1$), then p_j is an eigenvalue of $(\mathbf{P} + \mathbf{Q})$ of multiplicity at least $(m_j - 1)$ and at most $(m_j + 1)$.

The discussion in Appendix B shows that Theorem VIII is representative of more general results that can be proved by remarkably simple arguments centering about Lemma IV. Eigenvalue inequalities of the type discussed in the Appendix are ordinarily deduced from the extremal properties of eigenvalues.⁸

APPENDIX A.

Proof of Lemma I

Since $\mathbf{X}_1\mathbf{X}_2 = \mathbf{X}$, Lemma I is a direct consequence of the following result.

Lemma A: If \mathbf{D} and \mathbf{E} respectively are $n \times m$ and $m \times n$ matrices, det $[\mathbf{1}_n + \mathbf{DE}] = \det [\mathbf{1}_m + \mathbf{ED}]$.

We prove first that the lemma is true when \mathbf{D} and \mathbf{E} are square matrices. Let $\tilde{\mathbf{D}}$ and $\tilde{\mathbf{E}}$ be $p \times p$ matrices. Then, if $\tilde{\mathbf{D}}$ is nonsingular,

$$\det [\mathbf{1}_p + \tilde{\mathbf{D}}\tilde{\mathbf{E}}] = \det [\tilde{\mathbf{D}}^{-1}(\mathbf{1}_p + \tilde{\mathbf{D}}\tilde{\mathbf{E}})\tilde{\mathbf{D}}] = \det [\mathbf{1}_p + \tilde{\mathbf{E}}\tilde{\mathbf{D}}]. \quad (27)$$

If $\tilde{\mathbf{D}}$ is singular, it has a zero characteristic root, and hence there exists a positive number σ_0 such that $\tilde{\mathbf{D}} + \sigma\mathbf{1}_p$ is nonsingular for all real σ satisfying $0 < |\sigma| < \sigma_0$. Thus when $0 < |\sigma| < \sigma_0$,

$$\det [\mathbf{1}_p + (\tilde{\mathbf{D}} + \sigma\mathbf{1}_p)\tilde{\mathbf{E}}] = \det [\mathbf{1}_p + \tilde{\mathbf{E}}(\tilde{\mathbf{D}} + \sigma\mathbf{1}_p)]. \quad (28)$$

Both sides of (28) are polynomials in σ of at most degree p . Furthermore these polynomials must be identical since they agree throughout the real interval $(0, \sigma_0)$. Therefore (28) is valid when $\sigma = 0$.

Consider now the cases in which \mathbf{D} and \mathbf{E} are not square. Let $p = m + n$,

$$\tilde{\mathbf{D}} = \begin{matrix} & \begin{matrix} m & n \end{matrix} \\ \begin{bmatrix} \mathbf{D} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{bmatrix} & \begin{matrix} n \\ m \end{matrix} \end{matrix}, \quad \tilde{\mathbf{E}} = \begin{matrix} \begin{matrix} n & m \end{matrix} \\ \begin{bmatrix} \mathbf{E} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{bmatrix} & \begin{matrix} m \\ n \end{matrix} \end{matrix};$$

and let the symbol $\dot{+}$ denote a direct sum of matrices. Observe that $\det [\mathbf{1}_p + \tilde{\mathbf{D}}\tilde{\mathbf{E}}] = \det [(\mathbf{1}_n + \mathbf{DE}) \dot{+} \mathbf{1}_m] = \det [\mathbf{1}_n + \mathbf{DE}]$, and that $\det [\mathbf{1}_p + \tilde{\mathbf{E}}\tilde{\mathbf{D}}] = \det [(\mathbf{1}_m + \mathbf{ED}) \dot{+} \mathbf{1}_n] = \det [\mathbf{1}_m + \mathbf{ED}]$.

This proves the lemma.

APPENDIX B.

On the Eigenvalues of a Sum of Matrices: an Application of Lemma IV

In the following discussion \mathbf{M}^* and $\rho(\mathbf{M})$ respectively denote the

complex-conjugate transpose and the rank of an arbitrary matrix \mathbf{M} .

Our principal result is

Theorem A: Let $\{a_1, a_2, \dots, a_n\}$ denote the set of eigenvalues of the hermitian matrix \mathbf{A} , where the first p eigenvalues vanish if \mathbf{A} is of nullity p , and

$$-(a_{p+1})^{-1} \leq -(a_{p+2})^{-1} \leq \dots \leq -(a_n)^{-1}.$$

Let \mathbf{B} denote a nonnegative definite hermitian matrix of unit rank and order n ; and let

$$k = \lim_{s \rightarrow \infty} \frac{\det [\mathbf{1}_n + s(\mathbf{A} + \mathbf{B})]}{\det [\mathbf{1}_n + s\mathbf{A}]}.$$

Then $-\infty < k \leq \infty$ and the set of eigenvalues of $(\mathbf{A} + \mathbf{B})$ can be written as $\{c_1, c_2, \dots, c_n\}$ where

1. *if $0 < k < \infty$, the first p eigenvalues vanish and $-(a_{p+1})^{-1} \leq -(c_{p+1})^{-1} \leq -(a_{p+2})^{-1} \leq -(c_{p+2})^{-1} \leq \dots \leq -(a_n)^{-1} \leq -(c_n)^{-1}$*
2. *if $-\infty < k < 0$, the first p eigenvalues vanish and $-(c_{p+1})^{-1} \leq -(a_{p+1})^{-1} \leq -(c_{p+2})^{-1} \leq -(a_{p+2})^{-1} \leq \dots \leq -(c_n)^{-1} \leq -(a_n)^{-1}$*
3. *if $k = 0$, the first $(p + 1)$ eigenvalues vanish and $-(a_{p+1})^{-1} \leq -(c_{p+2})^{-1} \leq -(a_{p+2})^{-1} \leq -(c_{p+3})^{-1} \leq \dots \leq -(c_n)^{-1} \leq -(a_n)^{-1}$*
4. *if $k = \infty$, the first $(p - 1)$ eigenvalues vanish and $-(c_p)^{-1} \leq -(a_{p+1})^{-1} \leq -(c_{p+1})^{-1} \leq -(a_{p+2})^{-1} \leq \dots \leq -(a_n)^{-1} \leq -(c_n)^{-1}$.*

Proof:

The statements relating to the number of vanishing eigenvalues of $(\mathbf{A} + \mathbf{B})$ are a direct consequence of (i) $\rho(\mathbf{A}) = \rho(\mathbf{A} + \mathbf{B})$ when $0 < |k| < \infty$, (ii) $\rho(\mathbf{A}) = 1 + \rho(\mathbf{A} + \mathbf{B})$ when $k = 0$, and (iii) $\rho(\mathbf{A}) = \rho(\mathbf{A} + \mathbf{B}) - 1$ when $k = \infty$.

Consider now the real rational function in s :

$$y(s) = \frac{\det [\mathbf{1}_n + s(\mathbf{A} + \mathbf{B})]}{\det [\mathbf{1}_n + s\mathbf{A}]}$$

It is well-known that there exist two unitary matrices \mathbf{P} and \mathbf{Q} such that $\mathbf{PAP}^{-1} = \text{diag } [a_1, a_2, \dots, a_n]$ and $\mathbf{Q}(\mathbf{A} + \mathbf{B})\mathbf{Q}^{-1} = \text{diag } [c_1, c_2, \dots, c_n]$ where the a_i and c_i are real. Hence

$$y(s) = \prod_{i=1}^n (1 + sc_i)(1 + sa_i)^{-1}.$$

Consider the evaluation of $y(s)$ in accordance with Lemma IV. Here \mathbf{B} can be written as \mathbf{DD}^* where \mathbf{D} is an n -vector. Thus

$$\begin{aligned}
 y(s) &= 1 + s\mathbf{D}^*[\mathbf{I}_n + s\mathbf{A}]^{-1}\mathbf{D} \\
 &= 1 + s\mathbf{D}^*\mathbf{P}^{-1}[\text{diag}(1 + sa_1, 1 + sa_2, \dots, 1 + sa_n)]^{-1}\mathbf{PD} \\
 &= 1 + s\mathbf{H}^*\text{diag}[(1 + sa_1)^{-1}, (1 + sa_2)^{-1}, \dots, (1 + sa_n)^{-1}]\mathbf{H} \\
 &= 1 + \sum_{i=1}^n \frac{s |h_i|^2}{1 + sa_i} \tag{29}
 \end{aligned}$$

where $\mathbf{H} = \mathbf{PD} = [h_1, h_2, \dots, h_n]^t$. The theorem is obviously true when $\sum_{i=1}^n |h_i|^2 = 0$. Assume therefore that $\sum_{i=1}^n |h_i|^2 > 0$.

Observe that $y'(s) > 0$ for all real s such that $|y(s)| < \infty$, and that the poles of $y(s)$ are simple. Thus the interval on the real-axis of the s -plane between two adjacent poles contains one and only one zero of $y(s)$. Note that when $\lim_{s \rightarrow \infty} y(s) = k > 0$, the right-most critical point of $y(s)$ is a zero and that this critical point is a pole if $-\infty < k \leq 0$. Similarly the left-most critical point is a zero if $k = \infty$ or $-\infty < k < 0$; it is a pole if $0 \leq k < \infty$. The inequalities stated in the theorem follow directly from these observations. The equal signs make provision for possible coincident nonzero eigenvalues of \mathbf{A} and $(\mathbf{A} + \mathbf{B})$. Note that (29) implies that $-\infty < k \leq \infty$.

The following corollary of Theorem A appears to be useful in the study of linear dynamical systems.

Corollary A: Let $\tilde{\mathbf{A}}$ and $\tilde{\mathbf{B}}$ denote two hermitian matrices of order n , with $\tilde{\mathbf{A}}$ nonnegative definite. Let $\{\tilde{a}_1, \tilde{a}_2, \dots, \tilde{a}_n\}$ in which $\tilde{a}_1 \leq \tilde{a}_2 \leq \dots \leq \tilde{a}_n$ denote the set of eigenvalues of $\tilde{\mathbf{A}}$; and let $\{\tilde{c}_1, \tilde{c}_2, \dots, \tilde{c}_n\}$ in which $\tilde{c}_1 \leq \tilde{c}_2 \leq \dots \leq \tilde{c}_n$ denote the set of eigenvalues of $(\tilde{\mathbf{A}} + \tilde{\mathbf{B}})$. Then

$$\begin{aligned}
 \tilde{a}_{i-\beta} &\leq \tilde{c}_i & [1 + \beta \leq i \leq n] \\
 \tilde{c}_i &\leq \tilde{a}_{i+\alpha} & [1 \leq i \leq (n - \alpha)]
 \end{aligned}$$

where $\alpha = \frac{1}{2}(\tilde{r} + \tilde{s})$, $\beta = \frac{1}{2}(\tilde{r} - \tilde{s})$ in which \tilde{r} and \tilde{s} respectively are the rank and signature of $\tilde{\mathbf{B}}$.

Proof:

Consider Theorem A and assume that \mathbf{A} is nonnegative definite. Then either $k > 0$ or $k = \infty$ [i.e., either $\rho(\mathbf{A} + \mathbf{B}) = \rho(\mathbf{A})$ or $\rho(\mathbf{A} + \mathbf{B}) = \rho(\mathbf{A}) + 1$]. Thus in either case, since the c_i and a_i are nonnegative here,

$$a_1 \leq c_1 \leq a_2 \leq c_2 \dots \leq a_n \leq c_n \tag{30}$$

We may express $\tilde{\mathbf{B}}$ as $\sum_{i=1}^{\alpha} \tilde{\mathbf{B}}_i - \sum_{i=\alpha+1}^r \tilde{\mathbf{B}}_i$ where the $\tilde{\mathbf{B}}_i$ are rank 1 nonnegative definite hermitian matrices. It is certainly true that \tilde{c}_i does not exceed the corresponding eigenvalue of $\tilde{\mathbf{A}} + \sum_{i=1}^{\alpha} \tilde{\mathbf{B}}_i$ and is not less

than the corresponding eigenvalue of $\tilde{\mathbf{A}} - \sum_{i=\alpha+1}^r \tilde{\mathbf{B}}_i$. Hence by an α -fold application of inequalities of the type (30) we obtain the upper bound on \tilde{c}_i stated in the corollary. If \mathbf{B} in Theorem A were a nonpositive definite matrix the inequalities in (30) would be reversed, and hence a β -fold application of the inequalities suffices to establish the lower bound on \tilde{c}_i .

Our final result relates to eigenvalue multiplicities.

Theorem B: Let $\tilde{\mathbf{A}}$ and $\tilde{\mathbf{B}}$ be matrices of order n . Let $\tilde{\mathbf{A}}$ be similar to a diagonal matrix and let the rank of $\tilde{\mathbf{B}}$ be unity. Then if \tilde{a} is an eigenvalue of $\tilde{\mathbf{A}}$ of multiplicity m ($m \geq 1$), \tilde{a} is an eigenvalue of $(\tilde{\mathbf{A}} + \tilde{\mathbf{B}})$ of multiplicity at least $(m - 1)$. Further, if in addition $\tilde{\mathbf{A}}$ and $\tilde{\mathbf{B}}$ are hermitian, \tilde{a} is an eigenvalue of $(\tilde{\mathbf{A}} + \tilde{\mathbf{B}})$ of multiplicity at most $(m + 1)$.

Proof:

Consider

$$\tilde{y}(s) = \frac{\det[\mathbf{1}_n + s(\tilde{\mathbf{A}} + \tilde{\mathbf{B}})]}{\det[\mathbf{1}_n + s\tilde{\mathbf{A}}]}$$

Since $\tilde{\mathbf{B}}$ is of unit rank and $\tilde{\mathbf{A}}$ is similar to a diagonal matrix, the lemma can be used, as in the proof of Theorem A, to show that $\tilde{y}(s)$ has only simple poles. This proves the first part of the theorem which is essentially equivalent to the statement that if the dimensionality of the null-space of $[\tilde{a}\mathbf{1}_n - \tilde{\mathbf{A}}]$ is m , then the dimensionality of the null-space of $[\tilde{a}\mathbf{1}_n - (\tilde{\mathbf{A}} + \tilde{\mathbf{B}})]$ is not less than $(m - 1)$ [assuming that the Jordan form of $\tilde{\mathbf{A}}$ is diagonal and $\tilde{\mathbf{B}}$ is of unit rank]. It is not difficult to produce proofs of the result which are based on this interpretation.

If $\tilde{\mathbf{A}}$ and $\tilde{\mathbf{B}}$ are hermitian, $\tilde{y}(s)$ is either positive for all real s such that $|\tilde{y}(s)| < \infty$ or is negative for all such s (we assume that $\tilde{y}(s) \neq 1$), since $\tilde{\mathbf{B}}$ must be either nonnegative definite or nonpositive definite. Thus $\tilde{y}(s)$ can have only simple zeros, from which the second part of the theorem follows at once.

REFERENCES

1. Black, H. S., U. S. Patent No. 2,102,671.
2. Nyquist, H., Regeneration Theory, B.S.T.J., **11**, January 1932, pp. 126-147.
3. Blackman, R. B., Effect of Feedback on Impedance, BSTJ, **22**, October 1943, pp. 268-277.
4. Bode, H. W., *Network Analysis and Feedback Amplifier Design*, New York, D. Van Nostrand, 1945.
5. Mason, S. J., Feedback Theory — Some Properties of Signal Flow Graphs, Proc. IRE, **41**, September 1953, pp. 1144-1156.
6. Truxal, J. G., *Control System Synthesis*, New York, McGraw-Hill, 1955.
7. Mulligan, J. H. Jr., Signal Transmission in Non-reciprocal Systems, MRI Symposium Proceedings, Polytechnic Institute of Brooklyn, Vol. X, 1960, pp. 125-153.
8. Riesz, F. and Sz-Nagy, B., *Functional Analysis*, New York, Frederick Ungar Publishing Co., 1955, pp. 237-239.

On Overflow Processes of Trunk Groups with Poisson Inputs and Exponential Service Times

By A. DESCLOUX

(Manuscript received November 30, 1962)

Recurrence formulas for the Laplace transforms and the moments of the interoverflow distributions are obtained under the assumptions that the traffic offered is random (Poissonian) and that the service times are independent of each other and have a common negative exponential law. Under the same assumptions, it is also shown that the distribution of the nonbusy period of a group of c trunks is identical to the interoverflow distribution of a group of $c - 1$ trunks and that the distribution of the number of consecutive successful calls is essentially a mixture of geometric distributions. Processes obtained by superposing two or more overflow processes from independent trunk groups are not of the renewal type because interoverflow intervals are no longer independent. It is shown here that the correlation between two consecutive interoverflow intervals of a composite overflow process is always positive.

I. INTRODUCTION

When telephone networks are engineered, the loads offered to the alternate routes cannot usually be assumed to be random (Poissonian) and, in these cases, it is of considerable practical interest to determine the characteristics of the traffic overflowing the trunk groups. In this paper we shall be mainly concerned with the interoverflow distribution, the term used here to designate the distribution of the time intervals separating consecutive epochs at which calls find all trunks busy (overflow). We shall first show that the distribution of the nonbusy periods of a group of c trunks is identical to the interoverflow distribution of a group of $c - 1$ trunks and then obtain new recurrence formulas for the Laplace transforms of the interoverflow distribution of a single trunk group under the assumptions that: (a) the load submitted to the group is random; (b) the service times are independent of each other and are

all distributed according to the same negative exponential law; and (c) the requests which are placed when all the trunks are busy are either canceled or sent via some alternate route. As we shall see, these formulas are much simpler than the expressions obtained by C. Palm (cf Ref. 1, pp. 25-26, and Ref. 2, pp. 36-40) and are well suited to the computation of the moments. Then, under the same three assumptions, we shall also obtain the generating function of the probability distribution of the number of consecutive successful calls or, in other words, of the number of calls which are placed during a time interval whose end points coincide with two consecutive overflows. From the form of this generating function we can then infer that the distribution of the number of consecutive successful calls is essentially a mixture of geometric distributions.

In the remaining part of the paper we consider processes obtained by superposing (pooling) overflow or, more generally, renewal processes. In particular, it is shown that for processes obtained by superposition of two or more overflow processes, the covariance between the lengths of two intervals determined by three consecutive overflows is always positive.

II. RECURRENCE FORMULAS FOR THE MOMENTS OF THE INTEROVERFLOW DISTRIBUTION.

Consider a group of c trunks and assume that:

- i.* the calls are placed at random (Poisson input);
- ii.* the service times are independent of each other and of the state of the system and are distributed according to the negative exponential law with mean 1; and
- iii.* calls arriving when all trunks are busy do not wait but are either canceled or overflow to some other route (loss system).

Under these assumptions, the epochs at which overflows occur constitute a renewal process. We note also that if at some instant t the c trunks are busy, then the distribution of the time that elapses between t and the first overflow occurring after t is the same as the distribution of the intervals separating successive overflows. The cumulative distribution, $F_c(\cdot)$, of these intervals will be called the interoverflow distribution.

Before proceeding, we recall the following definitions of the busy and nonbusy periods: a busy period is a time interval during which all the trunks are continuously busy, and a nonbusy period is a time interval separating two consecutive busy periods. Under the present assumptions, the distribution of the busy periods is the negative exponential law with mean $1/c$.

Now, let $\gamma_c(\cdot)$ be the Laplace transform of the derivative of $F_c(\cdot)$. Then, as shown in Ref. 1, pp. 25-26, and Ref. 2, pp. 36-40,

$$\gamma_c(s) = D_c(s)/D_{c+1}(s) \tag{1}$$

with

$$D_n(s) = 1 + \sum_{j=1}^n \binom{n}{j} a^{-j} s(s+1) \cdots (s+j-1)$$

and a the demand rate. (Note that $D_n(s) = c_n(-s, a)$ with $c_n(\cdot, a)$ the Poisson-Charlier polynomial of degree n and parameter a .)

The roots r_1, r_2, \dots, r_{c+1} of $D_{c+1}(\cdot)$ are all negative and distinct so that:

$$F_c(t) = \sum_{i=1}^{c+1} k_i [1 - \exp(-s_i t)], \quad r_i = -s_i, \quad i = 1, \dots, c+1$$

where the coefficients $k_i, i = 1, \dots, c+1$ are given by the relations:

$$k_i s_i \prod_{j \neq i} (s_j - s_i) = a^{c+1} D_c(-s_i), \quad i = 1, 2, \dots, c+1.$$

We shall now derive the distribution of the nonbusy period. To this end, let us consider a busy period starting at time 0, say, and let T be the epoch at which the next following overflow occurs. We distinguish now between two cases:

1. At least one call is placed during the busy period under consideration. In this instance, the conditional density function, $f_1(\cdot)$, of T has the following expression:

$$f_1(t) = (c + a) \exp [-(c + a)t].$$

2. No call is placed during the busy period under consideration. In this case, the interval T is made up of three independent subintervals, namely, the busy period, the nonbusy period that follows it, and the period that elapses from the time at which all the trunks are made busy again to the occurrence of the first overflow. Since the distribution of this last subinterval is also $F_c(\cdot)$, the conditional density function, $f_2(\cdot)$, of T is then:

$$f_2(t) = (c + a) \int_0^t \int_0^v \exp [-(c + a)u] h_c(v - u) F_c'(t - v) du \cdot dv$$

where $h_c(\cdot)$ is the density function of the nonbusy period and $F_c'(\cdot)$ is the derivative of $F_c(\cdot)$.

Cases (1) and (2) occur with probabilities respectively equal to

$a/(c + a)$ and $c/(c + a)$ and we have therefore:

$$F'_c(t) = a \exp [-(c + a)t] + c \int_0^t \int_0^u \cdot \exp [-(c + a)u] h_c(v - u) F'_c(t - v) du \cdot dv. \quad (2)$$

Taking the Laplace transform on both sides of (2) yields:

$$\gamma_c(s) = \frac{a}{(c + a + s)} + \frac{c}{(c + a + s)} \cdot \gamma_c(s) \cdot \varphi_c(s) \quad (3)$$

with $\varphi_c(\cdot)$ the transform of $h_c(\cdot)$. Hence, solving for $\varphi_c(\cdot)$, we find

$$\begin{aligned} \varphi_c(s) &= (c + a + s)/c - (a/c)[\gamma_c(s)]^{-1} \\ &= \frac{(c + a + s)D_c(s) - aD_{c+1}(s)}{cD_c(s)}. \end{aligned}$$

Since (cf Ref. 2, p. 38 and p. 83)

$$aD_{c+1}(s) = (c + a + s)D_c(s) - cD_{c-1}(s)$$

it follows that:

$$\varphi_c(s) = \frac{D_{c-1}(s)}{D_c(s)}$$

which is the same as (1) with c replaced by $c - 1$. This shows that the nonbusy period distribution of a group of c trunks is identical to the interoverflow distribution of a group of $c - 1$ trunks with the same demand rate and average service time. (Note also that, under the present circumstances, the nonbusy period distribution remains unchanged if the calls finding all trunks occupied are allowed to wait.)

We have, therefore:

$$\varphi_c(s) = \gamma_{c-1}(s)$$

and (3) can be rewritten as follows:

$$\gamma_c(s) = a[c + a + s - c\gamma_{c-1}(s)]^{-1}. \quad (4)$$

It is interesting to note that this recurrence formula is considerably simpler than the one obtained by Palm,¹ namely,

$$\gamma_c(s) = \gamma_{c-1}(s + 1)[1 - \gamma_{c-1}(s) + \gamma_{c-1}(s + 1)]^{-1}. \quad (5)$$

The latter, however⁴, is valid for arbitrary recurrent input while (4) was obtained under the stronger assumption that the input is Poissonian. As one may expect, (4) can easily be obtained directly from (5). Indeed,

the recurrences (cf Ref. 2, p. 38 and p. 83)

$$aD_{c+1}(s) = (c + a + s)D_c(s) - cD_{c-1}(s)$$

and

$$D_{c+1}(s) = D_c(s) + [D_1(s) - 1]D_c(s + 1)$$

imply that

$$\begin{aligned} \gamma_{c-1}^{-1}(s + 1) &= \frac{D_c(s + 1)}{D_{c-1}(s + 1)} = \frac{D_{c+1}(s) - D_c(s)}{D_c(s) - D_{c-1}(s)} \\ &= \frac{1}{a} \frac{(c + a + s)D_c(s) - cD_{c-1}(s)}{D_c(s) - D_{c-1}(s)} - \frac{D_c(s)}{D_c(s) - D_{c-1}(s)} \\ &= \frac{1}{a} \frac{(c + a + s) - c\gamma_{c-1}(s)}{1 - \gamma_{c-1}(s)} - \frac{1}{1 - \gamma_{c-1}(s)}. \end{aligned}$$

Equation (4) then follows by substituting this expression in

$$\frac{1}{\gamma_c(s)} = 1 + \frac{1 - \gamma_{c-1}(s)}{\gamma_{c-1}(s + 1)}$$

which is Palm's recurrence in the reciprocal form.

Equation (4) can be used to obtain recurrences for the moments of the interoverflow distribution. Indeed, writing $\mu_n(c)$ for the n th moment of $F_c(\cdot)$, we find upon taking the n th derivative on both sides of (4):

$$\begin{aligned} \left. \frac{d^n}{ds^n} \gamma_c(s) \right|_{s=0} &= (-1)^n \mu_n(c) = a \frac{d^n}{ds^n} [c + a + s - c\gamma_{c-1}(s)]_{s=0}^{-1} \\ &= Y_n(fg_1, \dots, fg_n) \end{aligned}$$

where $Y_n(fg_1, \dots, fg_n)$ is a multivariable Bell polynomial (cf Ref. 3, p. 34-35 and p. 49) with:

$$\begin{aligned} f_k &= (-1)^k k! a^{-k} \\ g_1 &= 1 + c\mu_1(c - 1) \\ g_k &= (-1)^{k+1} c\mu_k(c - 1), \quad k > 1. \end{aligned}$$

Since

$$\mu_n(0) = n! a^{-n}, \quad n = 0, 1, 2, \dots$$

these relations can be used to compute the moments of the interoverflow distribution recurrently. In particular, we have:

$$a\mu_1(c) = 1 + c\mu_1(c - 1) = E_{1,c}^{-1}(a)$$

where $E_{1,c}(a)$, which is known as the first Erlang loss function, is the probability that a call is placed when all trunks are busy and is therefore cleared from the system. For $n = 2, 3$ and 4 we have the following recurrences:

$$\begin{aligned} a\mu_2(c) &= 2a\mu_1^2(c) + c\mu_2(c-1) \\ a\mu_3(c) &= 6a\mu_1^3(c) + 6c\mu_1(c)\mu_2(c-1) + c\mu_3(c-1) \\ a^2\mu_4(c) &= 24a^2\mu_1^4(c) + 36ac\mu_1^2(c)\mu_2(c-1) + 6c^2\mu_2^2(c-1) \\ &\quad + 8ac\mu_1(c)\mu_3(c-1) + ac\mu_4(c-1). \end{aligned}$$

Finally we note that repeated use of the first of these relations yields the following expression for the second moment:

$$\mu_2(c) = 2 \sum_{n=0}^c \frac{(c)_n}{a^n} \mu_1^2(c-n)$$

with $(c)_0 = 1$ and $(c)_n = c(c-1) \cdots (c-n+1)$.

III. DISTRIBUTION OF THE NUMBER OF CONSECUTIVE SUCCESSFUL CALLS

In this section we consider a group of c indexed trunks with calls always assigned to the free trunk having the lowest index. The calls which find the first m trunks busy will be referred to as m -overflows, and those which are placed when at least one of the first m trunks is free will be said to be m -successful. (c -successful calls are simply said to be successful.)

We shall designate by $F_m(\cdot)$ the cumulative distribution of the time interval separating two consecutive m -overflows and by $P_m(t, n)$ the probability that exactly $n-1$ calls are m -successful during a time interval of length t whose end points coincide with two consecutive epochs at which m -overflows occur.

Taking the average service time as unity, we have then the following recurrence:

$$\begin{aligned} P_c(t, n) dF_c(t) &= e^{-t} P_{c-1}(t, n) dF_{c-1}(t) \\ &\quad + \sum_{k=1}^{n-1} \int_0^t (1 - e^{-u}) P_c(t-u, n-k) P_{c-1}(u, k) dF_c(t-u) dF_{c-1}(u). \end{aligned} \quad (6)$$

Indeed, let us assume that a c -overflow occurs at time 0 . Then the event "the first c -overflow after time 0 occurs in the interval $(t, t + \Delta t)$ and there are $n-1$ successful calls during $(0, t)$ " can be split as follows:

i. The first c -overflow after time 0 occurs during $(t, t + \Delta t)$; the call

being served by the c th trunk at time 0 does not terminate before t ; and the number of $(c - 1)$ -successful calls is equal to $n - 1$. The probability of this event is equal to

$$e^{-t}P_{c-1}(t,n)\Delta F_{c-1}(t) + o(\Delta t).$$

ii. The first $(c - 1)$ -overflow after time 0 occurs during $(u, u + \Delta u)$, $u < t$; the first c -overflow after time 0 occurs during $(t, t + \Delta t)$; the call being served by the c th trunk at time 0 terminates before u ; and the number of $(c - 1)$ -successful calls during $(0, u)$ is equal to $k - 1$, while the number of c -successful calls during (u, t) is equal to $n - k - 1$, $k = 1, \dots, n - 1$. The probability of this event is, in first approximation:

$$(1 - e^{-u})P_c(t - u, n - k)P_{c-1}(u, k)\Delta F_c(t - u)\Delta F_{c-1}(u).$$

Equation (6) is then obtained by summing up these probabilities and then passing to the limit $(\Delta u \rightarrow 0, \Delta t \rightarrow 0)$ and integrating with respect to u $(0 \leq u \leq t)$.

Now write:

$$\lambda_c(x, w) = \sum_{n=1}^{\infty} x^n \int_0^{\infty} e^{-wt} P_c(t, n) dF_c(t).$$

Then (6) implies:

$$\lambda_c(x, w) = \frac{\lambda_{c-1}(x, w + 1)}{1 - \lambda_{c-1}(x, w) + \lambda_{c-1}(x, w + 1)} \tag{7}$$

which is of the same form as Palm's recurrence (5). (Note that while (5) holds for arbitrary recurrent input, (7) is valid only when the input is Poissonian.)

Clearly, $\lambda_c(x, w)$ can be written as a ratio†:

$$\lambda_c(x, w) = \frac{D_c(x, w)}{D_{c+1}(x, w)}. \tag{8}$$

Substituting this expression in (7), we find that

$$\begin{aligned} \frac{D_{c+1}(x, w) - D_c(x, w)}{D_c(x, w + 1)} &= \frac{D_c(x, w) - D_{c-1}(x, w)}{D_{c-1}(x, w + 1)} = \dots \\ &= \frac{D_1(x, w) - D_0(x, w)}{D_0(x, w + 1)}. \end{aligned} \tag{9}$$

† The method used here to solve (7) is formally identical to the one used in Ref. 2 to solve Palm's recurrence (5).

Setting $D_0(x, w) \equiv 1$, which is not a restriction of the generality, we obtain, using (9):

$$D_{r+1}(x, w) = D_r(x, w) + [D_1(x, w) - 1]D_r(x, w + 1), \quad r \geq 1. \quad (10)$$

We note that:

$$\lambda_0(x, w) = \int_0^\infty e^{-wt} x dF_0(t) = ax \int_0^\infty \exp[-t(a+w)] dt = \frac{ax}{w+a}$$

so that: $D_1(x, w) = (w+a)/ax$.

Solving now (10) recurrently, we find:

$$D_m(x, w) = 1 + \sum_{j=1}^m \binom{m}{j} \left(\frac{1}{ax}\right)^j \prod_{k=1}^j [a(1-x) + w + (k-1)].$$

From this, it follows that the probability generating function, $H_c(\cdot)$, of the number of successful calls between two consecutive c -overflows, is

$$H_c(x) = \frac{\lambda_c(x, 0)}{x} = \frac{1}{x} \frac{D_c(x, 0)}{D_{c+1}(x, 0)}$$

with

$$D_m(x, 0) = 1 + \sum_{j=1}^m \binom{m}{j} \left(\frac{1}{ax}\right)^j \prod_{k=1}^j [a(1-x) + (k-1)].$$

These relations can be expressed in a simpler manner. Indeed (cf Ref. 2, p. 83):

$$axD_{m+1}(x, 0) = (m+a)D_m(x, 0) - mD_{m-1}(x, 0)$$

and, with the notation $N_m(x) = (ax)^m D_m(x, 0)$, we have, therefore:

$$H_c(x) = a \frac{N_c(x)}{N_{c+1}(x)}$$

where:

$$N_0(x) \equiv 1, \quad N_1(x) \equiv a$$

and

$$N_{m+1}(x) = (m+a)N_m(x) - axm N_{m-1}(x). \quad (11)$$

The polynomials $N_m(\cdot)$ have the following properties which are immediate consequences of (11):

- i. $N_m(x) > 0$ for $x \leq 0$, $m = 0, 1, 2, \dots$.
- ii. the degree $\nu = \nu(m)$ of $N_m(\cdot)$ is the integral part of $m/2$.
- iii. The coefficients of x^0, x^1, \dots, x^ν in $N_m(x)$ are alternately positive

and negative. Taking (ii) into account, it then follows, as x tends to ∞ , that $N_m(x)$ tends to ∞ if $\nu(m)$ is even and to $-\infty$ if $\nu(m)$ is odd.

Repeated application of (11) yields:

$$\begin{aligned} &(m + a - 2)N_{m+1}(x) + \{ax[m(m + a - 2) + (m + a)(m - 1)] \\ &\quad - (m + a)(m + a - 1)(m + a - 2)\}N_{m-1}(x) \quad (12) \\ &\quad + (m + a)(m - 1)(m - 2)(ax)^2N_{m-3}(x) = 0. \end{aligned}$$

Using (12) and the properties (i)-(iii) of the polynomials $N_m(\cdot)$, it is easily shown by consideration of the signs that $N_m(\cdot)$ has $\nu(m)$ distinct roots which are positive and separated by the $\nu(m) - 1$ roots of $N_{m-2}(\cdot)$, $m = 4, 5, \dots$. Let now $x_1 < x_2 < \dots < x_{\nu(c+1)}$ be the roots of $N_{c+1}(\cdot)$. Then the generating function $H_c(\cdot)$ is of the form:

$$H_c(x) = \gamma_0 + \sum_{i=1}^{\nu(c+1)} \gamma_i \left(1 - \frac{x}{x_i}\right)^{-1}$$

where $\gamma_0 = 0$ if c is odd and $\gamma_0 > 0$ if c is even, and where the constants $\gamma_i, i = 1, \dots, \nu(c + 1)$ are strictly positive.

Consequently,† the distribution of the number of successful calls is a mixture of $\nu(m)$ distinct geometric distributions when c is odd; when c is even, the distribution of the number of successful calls is a mixture of $\nu(m) + 1$ distributions, one of the latter being the distribution with probability mass 1 at the origin and the remaining ones being distinct geometric distributions.

Finally, we note that the recurrences

$$axD_{m+1}(x,w) = (m + w + a)D_m(x,w) - mD_{m-1}(x,w)$$

and

$$D_{m+1}(x,w) = D_m(x,w) + [D_1(x,w) - 1]D_m(x,w + 1)$$

allow us to write (7) in the simpler form

$$\lambda_c(x,w) = ax[c + a + w - c\lambda_{c-1}(x,w)].^{-1}$$

Hence, we also have

$$H_c(x) = a[c + a - cxH_{c-1}(x)]^{-1}$$

which may be used to compute the moments of the distribution of the number of consecutive successful calls recurrently.

† Since $H_c(\cdot)$ is analytic for $|x| < 1$, it follows that the roots $x_1, x_2, \dots, x_{\nu(c+1)}$ are all larger than 1.

IV. COVARIANCE BETWEEN INTEROVERFLOW INTERVALS

Under the assumptions made here, the overflow process of a single trunk group is of the renewal type. The processes obtained by superposing two or more such processes (called here composite overflow processes) do not, however, have this property because successive interoverflow intervals are no longer independent. We shall now prove that the correlation between two consecutive interoverflow intervals of a composite overflow process is always positive.

Let us consider n trunk groups G_1, G_2, \dots, G_n ($n \geq 2$), of sizes c_1, c_2, \dots, c_n , respectively, and let A_i be the (random) load offered to G_i . Let us now place ourselves at an overflow epoch and suppose that the overflow call in question comes from group G_1 . Let also U and V be the two consecutive interoverflow intervals separated by the overflow under consideration. Then, by an argument similar to one used by Cox and Smith,⁴ it follows that:

$$\Pr [U \geq u, V \geq v] = [1 - F_1(u)] \cdot [1 - F_1(v)] \prod_{i=2}^n \int_{u+v}^{\infty} \frac{1 - F_i(x)}{\mu_i} dx \quad (13)$$

where $F_i(\cdot)$ is the cumulative interoverflow distribution of G_i (considered by itself) and

$$\mu_i = \int_0^{\infty} x dF_i(x).$$

[Equation (13) implies that the process obtained by superposing continuous renewal processes is itself a renewal process if and only if all the distributions $F_i(\cdot)$ are negative-exponentials. An overflow process is therefore of the renewal type if and only if $n = 1$.]

Under the present assumptions we have:²

$$F_i(x) = \sum_{j=1}^{c_i+1} a_{ij} [1 - \exp(-s_{ij}x)], \quad x \geq 0 \quad (14)$$

$$F_i(x) = 0, \quad x < 0$$

where $a_{ij} > 0$, $s_{ij} > 0$, $j = 1, \dots, c_i + 1$, $i = 1, 2, \dots, n$, and

$$\sum_{j=1}^{c_i+1} a_{ij} = 1$$

$$\sum_{j=1}^{c_i+1} \frac{a_{ij}}{s_{ij}} = \mu_i.$$

Upon substituting (14) into (13) we find that:

$$\begin{aligned}
 & \Pr [U \geq u, V \geq v] \\
 &= \frac{1}{M_1} \left[\sum_{j=1}^{c_1+1} a_{1j} \exp(-s_{1j}u) \right] \left[\sum_{j=1}^{c_1+1} a_{1j} \exp(-s_{1j}v) \right] \\
 & \quad \cdot \prod_{i=2}^n \int_{u+v}^{\infty} \left[\sum_{j=1}^{c_i+1} a_{ij} \exp(-s_{ij}x) \right] dx \\
 &= \frac{1}{M_1} \left[\sum_{j=1}^{c_1+1} a_{1j} \exp(-s_{1j}u) \right] \left[\sum_{j=1}^{c_1+1} a_{1j} \exp(-s_{1j}v) \right] \tag{15} \\
 & \quad \cdot \prod_{i=2}^n \left[\sum_{j=1}^{c_i+1} \frac{a_{ij}}{s_{ij}} \exp[-s_{ij}(u+v)] \right] \\
 &= \frac{1}{M_1} \left[\sum_{j=1}^{c_1+1} a_{1j} \exp(-s_{1j}u) \right] \left[\sum_{j=1}^{c_1+1} a_{1j} \exp(-s_{1j}v) \right] \\
 & \quad \cdot \sum^* \frac{a_{2j_2} a_{3j_3} \cdots a_{nj_n}}{s_{2j_2} s_{3j_3} \cdots s_{nj_n}} \exp[-(s_{2j_2} + s_{3j_3} + \cdots + s_{nj_n})(u+v)]
 \end{aligned}$$

where the summation \sum^* is to be extended over all $(n - 1)$ -tuples (j_2, \dots, j_n) arising when multiplying out

$$\prod_{i=2}^n \left[\sum_{j=1}^{c_i+1} \frac{a_{ij}}{s_{ij}} \right].$$

From here on we shall use the letter J as a generic symbol for any one of these $(n - 1)$ -tuples.

Integrating now (15) with respect to u and v yields:

$$\int_0^{\infty} \int_0^{\infty} \Pr [U \geq u, V \geq v] du \cdot dv = \frac{1}{M_1} \sum^* m_1(J) R_1^2(J)$$

with:

$$m_1(J) = \frac{a_{2j_2} \cdots a_{nj_n}}{s_{2j_2} \cdots s_{nj_n}}, \quad J = (j_2, \dots, j_n)$$

and

$$R_1(J) = \sum_{j=1}^{c_1+1} \frac{a_{1j}}{s_{1j} + (s_{2j_2} + \cdots + s_{nj_n})}, \quad J = (j_2, \dots, j_n)$$

We also note that:

$$\begin{aligned}
 \mu_1 &= \int_0^{\infty} [1 - F_1(u)] \left[\prod_{i=2}^n \int_u^{\infty} \frac{1 - F_i(x)}{\mu_i} dx_i \right] du \\
 &= \frac{1}{M_1} \sum^* m_1(J) \cdot R_1(J).
 \end{aligned}$$

When the overflow call separating the two intervals U and V is from group G_i , $i \neq 1$, we define M_i , $m_i(J)$ and $R_i(J)$ to have the meaning corresponding to M_1 , $m_1(J)$ and $R_1(J)$, respectively. Let also P_i be the probability that the separating call is from group G_i . Then with μ denoting the average length between successive overflow calls and $\text{Cov}(U, V)$ standing for the covariance between U and V , we have:

$$\mu = \sum_{i=1}^n \frac{P_i}{M_i} \sum^* m_i(J) \cdot R_i(J)$$

and

$$\text{Cov}(U, V) = \sum_{i=1}^n \frac{P_i}{M_i} \sum^* m_i(J) R_i^2(J) - \mu^2.$$

We note now that:

$$\begin{aligned} \mu^2 &= \sum_{i=1}^n \frac{P_i}{M_i} \sum^* m_i(J) \left[\frac{R_i(J)}{\mu} - 1 \right]^2 \\ &= \sum_{i=1}^n \frac{P_i}{M_i} \sum^* m_i(J) R_i^2(J) - 2\mu^2 + \mu^2 \sum_{i=1}^n \frac{P_i}{M_i} \sum^* m_i(J). \end{aligned} \quad (16)$$

Therefore, since:

$$\sum^* m_i(J) = \prod_{\substack{k=i \\ k \neq i}}^n \int_0^{\infty} \sum_{j=1}^{c_k+1} [a_{kj} \exp(-s_{kj}x)] dx = M_i$$

we obtain, upon substituting this last expression in (16):

$$\text{Cov}(U, V) = \sum_{i=1}^n \frac{P_i}{M_i} \sum^* m_i(J) [R_i(J) - \mu]^2 \geq 0. \quad (17)$$

In order to complete the proof, we have to show that the equality sign in (17) cannot hold when $n > 1$. This, however, is an immediate consequence of the fact that the s_{ij} 's occurring in $F_i(\cdot)$ are then all distinct.

V. SOME ADDITIONAL PROPERTIES

It is stated in Ref. 4 that the processes obtained by superposing n identical renewal processes tend usually to renewal processes with negative-exponential distributions as n tends to infinity. To get an idea about the speed of that convergence, we shall examine how fast the correlation between two consecutive intervals, U , V , determined by three consecutive arrival epochs tends to zero (arrival epoch = epoch at which a renewal occurs in any one of the n processes).

Let $F(\cdot)$ be the cumulative distribution function of the intervals

separating pairs of consecutive renewal epochs from the same process. Under some mild restrictions, namely that $F(\cdot)$ and its first derivative are continuous in some closed interval $[0, \delta]$, $\delta > 0$, it can be shown that:

$$\lim_{n \rightarrow \infty} n^3 \text{Cov}(U, V) = \gamma \tag{18}$$

where γ , a constant, is strictly positive if $F'(0) > 1$, strictly negative if $F'(0) < 1$, and zero if $F'(0) = 1$. [Here $F'(0)$ has to be understood as being the right-hand derivative of $F(\cdot)$ at 0.] Further:

$$\text{Cov}(U, V) = O(n^{-4}), \quad n \rightarrow \infty, \quad \text{when } F'(0) = 1. \tag{19}$$

Let $\text{Var}(X)$ stand for the variance of X and $\rho(X, Y)$ for the coefficient of correlation of X and Y .

We have here $\text{Var}(U) = \text{Var}(V)$ and

$$\lim_{n \rightarrow \infty} n^2 \text{Var}(U) = \sigma$$

where σ is a strictly positive constant.

We have, therefore:

$$\lim_{n \rightarrow \infty} n \cdot \rho(U, V) = k \tag{20}$$

where k , a constant, is strictly positive if $F'(0) > 1$, strictly negative if $F'(0) < 1$, and zero if $F'(0) = 1$. Further:

$$\rho(U, V) = O(n^{-2}), \quad n \rightarrow \infty, \quad \text{when } F'(0) = 1. \tag{21}$$

We proceed now with the proof of (18) and (19). Assuming that $EU = EV = 1$, we have:

$$\begin{aligned} \text{Cov}(U, V) + n^{-2} &= \int_0^\infty \int_0^\infty [1 - F(u)][1 - F(v)] \exp[(n - 1)H(u + v)] \\ &\quad \cdot du \cdot dv \end{aligned}$$

with

$$H(x) = \log \int_x^\infty [1 - F(t)] dt.$$

Let now $u = (y + z)/2$, $v = (z - y)/2$. Then:

$$\begin{aligned} \text{Cov}(U, V) + n^{-2} &= \int_0^\infty \exp[(n - 1)H(z)] \int_0^z \left[1 - F\left(\frac{z + y}{2}\right)\right] \left[1 - F\left(\frac{z - y}{2}\right)\right] \\ &\quad \cdot dz \cdot dy. \end{aligned}$$

We note, however, that:

$$x \log \int_{x^{-(k+\epsilon)}}^{\infty} [1 - F(t)] dt = - |0(x^{k-\epsilon})|, \quad 0 \leq \epsilon < \frac{1}{2}, \quad x \rightarrow \infty$$

so that:

$$\begin{aligned} \text{Cov}(U, V) + n^{-2} &\sim \int_0^{(n-1)^{-(k+\epsilon)}} \exp [(n-1)H(z)] \\ &\cdot \int_0^z \left[1 - F\left(\frac{z+y}{2}\right) \right] \left[1 - F\left(\frac{z-y}{2}\right) \right] \cdot dz \cdot dy \\ &= \int_0^{(n-1)^{-(k+\epsilon)}} \exp [(n-1)H(z)] [z - F'(\theta z)z^2 + O(z^3)] \cdot dz \\ &\qquad\qquad\qquad 0 \leq \theta \leq 1, \quad 0 \leq \epsilon < \frac{1}{2}. \end{aligned}$$

Upon expanding $\exp (n-1)H(z)$ and making the substitution $t = (n-1)z$, we find:

$$\begin{aligned} \text{Cov}(U, V) + n^{-2} &\sim \frac{1}{(n-1)^2} \int_0^{(n-1)^{k-\epsilon}} e^{-t} \left\{ 1 + \frac{F'(\theta' t/n-1) - 1}{2} \right. \\ &\cdot \left. \left[\frac{t^2}{n-1} + t^3 O(n^{-2}) \right] \right\} \left[t - F'(\theta t/n-1) \frac{t^2}{n-1} + t^3 O(n^{-2}) \right] \cdot dt \end{aligned}$$

where $0 \leq \theta' \leq 1$.

We can therefore conclude that:

$$\begin{aligned} \text{Cov}(U, V) &\sim \frac{1}{(n-1)^2} - \frac{2F'(0)}{(n-1)^3} + \frac{3[F'(0) - 1]}{(n-1)^3} - \frac{1}{n^2} \\ &+ O(n^{-4}) = \frac{F'(0) - 1}{(n-1)^3} + O(n^{-4}). \end{aligned}$$

This relation implies (18) and (19)

Finally, we shall state the following three properties:

i. If $F(\cdot)$ is the uniform distribution, then $\text{Cov}(U, V) < 0$ for all values of $n (\geq 2)$.

ii. For any given $n (\geq 2)$, it is always possible to choose $F(\cdot)$ in such a way that the correlation between U and V vanishes while the variables U and V are not independent.

iii. Let $n = 2$ and assume that:

$$F(x) = 0, \quad x < \alpha$$

$$F(x) = p, \quad \alpha \leq x < \beta$$

$$F(x) = 1, \quad \alpha \geq \beta$$

$$\text{where } \alpha p + \beta(1 - p) = 1.$$

Then $\text{Cov}(U, V) = 0$ if and only if

$$6p\alpha(1 - \alpha)^2 + 8(p\alpha)^2(1 - \alpha) + 2(p\alpha)^2(p\alpha - 1) - 1 = 0$$

When this condition is satisfied, $\rho(U, V) = 0$ although U and V are dependent.

VI. ACKNOWLEDGMENT

I wish to express my appreciation to John Riordan for many useful comments. I also wish to thank D. C. Boes and M. Segal for stimulating discussions.

REFERENCES

1. Palm, C., Intensitätsschwankungen im Fernsprechverkehr, Ericsson Technics, **44**, 1943, pp. 1-189.
2. Riordan, John, *Stochastic Service Systems*, John Wiley & Sons, Inc., New York, 1962.
3. Riordan, John, *An Introduction to Combinatorial Analysis*, John Wiley & Sons, Inc., New York, 1958.
4. Cox, D. R., Smith, W. L., On the Superposition of Renewal Processes, *Biometrika*, **41**, 1954, pp. 91-99.

Solar Cell Degradation under 1-Mev Electron Bombardment

By W. ROSENZWEIG, H. K. GUMMEL and F. M. SMITS

(Manuscript received July 23, 1962)

The effect of radiation damage on the important parameters of solar cells has been evaluated for groups of blue-sensitive n-on-p, normal p-on-n, and blue-sensitive p-on-n cells using 1-Mev electrons. The outer space short circuit current, maximum power, junction characteristic, and spectral response are presented quantitatively as a function of radiation flux along with the bulk minority-carrier diffusion length.

The rate of change of inverse squared diffusion length with flux is found to be 1.22×10^{-8} for the p-on-n cells, as compared to 1.7×10^{-10} for the n-on-p. The degradation of the spectral response is consistent with the measured diffusion length for both types of cells if one assumes a total effective front layer of 1 μ thickness for the p-on-n cells.

As a result of the less rapid degradation of their minority-carrier lifetime, the n-on-p cells exhibit a greater resistance to radiation than the p-on-n cells. Comparing only the two types of blue-sensitive cells, after prolonged bombardment, the flux ratios required to achieve equal values of short-circuit current and maximum power are 17 and 9.5 respectively.

I. INTRODUCTION

The high-energy electron and proton radiation in the Van Allen Belt produces defects in semiconductors which cause a reduction in solar cell power output. This raises immediately the problem of assessing the expected useful life of the power plant of a satellite that passes through this radiation belt. The life is the time for which the power plant is capable of delivering the electrical power necessary for successful operation of the satellite.

For an evaluation one has to determine the maximum power point under outer space solar illumination after various amounts of bombardment by protons of differing energies and by electrons of differing energies, with various thicknesses of shielding over the cells. Such

information may then be synthesized to yield the expected performance for a given Van Allen Belt spectrum under given shielding conditions. An attempt at a direct experimental investigation of these relations has the disadvantage of being exceedingly complex and involved. However, the complexity of this problem can be significantly reduced by recognizing that the solar cell performance is strongly dependent on the minority-carrier lifetime in the bulk of the material. This minority-carrier lifetime is ordinarily the most radiation-sensitive parameter in a semiconductor, and a solar cell will, practically speaking, be useless due to minority-carrier lifetime degradation before any of the other determining parameters are significantly affected by radiation.

The performance of a given type of solar cell after irradiation is determined almost entirely by the minority-carrier lifetime remaining after the irradiation, rather than by the type of radiation responsible for the lifetime degradation. The different radiations might cause degradations that lead to differences in temperature dependence of the solar cell performance. However, such differences, if present, can be expected to be unimportant for most practical applications.

These considerations imply that relative damage rates can be established for protons and electrons of various energies, from which the radiation damage due to the complex particle spectrum in the Van Allen Belt can be related to the radiation damage under any one monoenergetic radiation.

The problem can thus be divided into two broad studies. One is the detailed study of the solar cell characteristics, i.e. short-circuit current, maximum power, etc., as a function of exposure to a conveniently available monoenergetic radiation. The other is an investigation of the relative damage rates for the various radiations, from which a monoenergetic flux can be derived which is equivalent in radiation damage to the complex Van Allen spectrum.

With the over-all solar cell performance depending on a variety of process variables, a statistical experiment is advisable for the first part so that the effects of variations in solar cell structure are averaged.

For the investigation of the relative damage rates for the various radiations it is desirable to study the degradation of the minority-carrier lifetime, or of a quantity that is directly related to the minority-carrier lifetime, such as the minority-carrier diffusion length. Since this quantity is a bulk property of the material one is concerned with only one variable. In addition, since all the solar cell parameters are smoothly varying functions of the diffusion length, and the diffusion length varies with flux in a known way [see (1) below], one can perform useful inter-

polations and extrapolations on the values of the parameters to extremely low and high radiation fluxes.

This paper is devoted to the first part of such an over-all evaluation, i.e. the determination of the degradation of the important cell parameters on bombardment with radiation of one kind. It describes an experiment in which solar cell performance as a function of 1-Mev electron irradiation was investigated for samples large enough to give good statistical results. For comparison blue-sensitive n-on-p solar cells, blue-sensitive p-on-n, and normal p-on-n solar cells were studied.

First the details of the electron irradiations are given, followed by a general description of the performance tests at the various stages of the bombardment. The final sections summarize the results and give the conclusions of the experiment.

II. ELECTRON IRRADIATIONS AND DIFFUSION LENGTH MEASUREMENTS

The evaluation of solar cell performance was carried out with a 1-Mev electron Van de Graaff generator with the external beam scattered by a total of 5.5 mils of aluminum and 5 inches of air. Groups of 16 or more solar cells of the following types were used:

1. Blue-sensitive n-on-p cells produced by the Western Electric Company. These cells were randomly selected from a lot of 10,000 cells having an efficiency greater than 7.5 per cent under outer space light.

2. Normal p-on-n cells secured from a commercial source and rated as 14 per cent efficient under tungsten light.

3. Blue-sensitive p-on-n cells secured from a commercial source rated as 12 per cent efficient under outer space light.

To achieve uniform exposure for all cells, the solar cells were mounted near the perimeter of an aluminum disc which was continuously rotated during the irradiations in such a way that the cells passed through the center of the beam. The electron beam intensity was calibrated with a vacuum Faraday cup located on the beam axis at the same distance from the beam exit window as that at which the cells were irradiated. The irradiation duty cycle on the rotating wheel was measured by using a prebombarded solar cell as radiation intensity monitor¹ measuring the integrated exposure per turn referred to the center of the beam. The purpose of the prebombardment was to keep the cell diffusion length from changing during the measurement. The exposures were monitored with the Faraday cup positioned near the edge of the wheel off the beam axis.

The cells were irradiated in five steps to integrated fluxes of 1.8×10^{13} ,

9.0×10^{13} , 5.4×10^{14} , 2.7×10^{15} , and 1.8×10^{16} electrons/cm². Before the first and after each successive bombardment the cells were subjected to optical and electrical measurements by methods described in the next section. In addition to these measurements the minority-carrier diffusion length was also determined for each cell by measuring the short-circuit current response of the solar cells under a low intensity 1-Mev electron beam filtered by 12 mils of aluminum placed directly over the cell.² Under this condition excess carriers are generated uniformly throughout the bulk of the cell at a rate of 225 pairs/ μ per incident electron. Except at very short diffusion lengths as discussed below, the short-circuit current thus obtained is proportional to the diffusion length.

These diffusion length measurements permit a correlation between the results reported here and the radiation damage effects of other radiations.

III. MEASUREMENT OF SOLAR CELL PARAMETERS AND OUTER SPACE CALCULATIONS

The ultimate criterion for the performance of a solar power plant is the power it can deliver after a substantial exposure to radiation, regardless of the initial performance. This requires that the absolute performance of the solar cells under outer space illumination be studied. To permit the performance of such tests in the laboratory a solar cell test facility has been developed which is described in detail elsewhere.³

Prior to bombardments and after each bombardment step, all cells were tested in this facility. A set of control cells was kept with the cells under study and received the same handling except for the bombardments. These control cells were tested, interspersed among sample cells, each time the sample cells were tested. In this way, the stability of the test facility was monitored and corrections for small drifts could be applied. The largest corrections necessary amounted to about 1 per cent.

In the test facility the quantum efficiency, defined as electrons delivered into a short circuit per incident photon, is measured at eight wavelengths between 0.4 and 0.95 μ . Also, under a "white light" source of intensity approximately equal to outer space sunlight, short-circuit current, open-circuit voltage, and voltage across a 10-ohm load are measured. These quantities define the forward current-voltage characteristic of the diode. Some additional tests, such as a reverse leakage measurement, and the cell voltage for a 50-ma injected forward current in the dark are also performed.

The results of these measurements are automatically recorded on IBM cards and processed on an IBM 7090 computer.

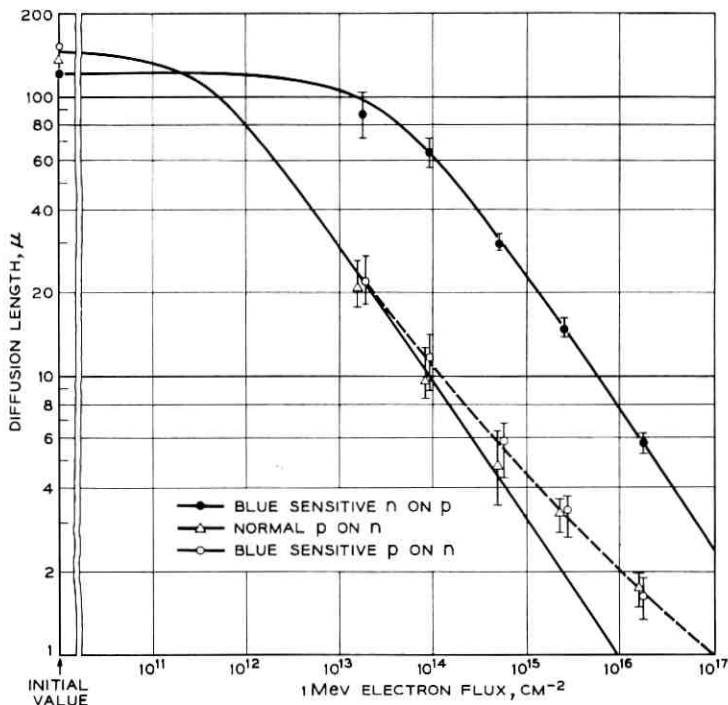


Fig. 1 — Diffusion length versus 1-Mev electron flux.

The outer space short-circuit current is computed from the quantum efficiencies. The remainder of the measurements are used to obtain the output characteristics of the cell corresponding to the outer space illumination and the maximum power point. Averages of the results are formed for each group. Generally, only these averages are quoted in the following section.

IV. EXPERIMENTAL RESULTS AND DISCUSSION

The results of the various measurements are presented in Figs. 1-11. A plot of the mean diffusion length in microns as a function of the integrated electron flux is given in Fig. 1. The limits placed around each point indicate not the experimental error but the rms deviation from the mean of the measured quantity for the given sample. It is to be noted that the deviation for the n-on-p cells becomes smaller as the flux is increased, indicating that the damage rates are quite similar for the cells in the entire sample.

The diffusion length is expected to decrease according to the equation

$$\frac{1}{L^2} = \frac{1}{L_0^2} + K\Phi \quad (1)$$

in which Φ is the bombardment flux in cm^{-2} , L is the diffusion length in cm for that flux, and L_0 is the initial diffusion length. Equation (1) is an expression of the fact that the recombination rate of the excess minority carriers at any point in the bombardment is proportional to the initial number of recombination centers present plus the number introduced during the bombardment, the latter being proportional to the flux. The solid curve passing through the points for the n-on-p cells was computed using (1) with $L_0 = 119 \mu$ and $K = 1.7 \times 10^{-10}$.

The experimental points for the p-on-n cells can be fitted, as shown in the dashed curve, by assuming a sensitive thickness consisting of the bulk diffusion length, as given by the solid line, and a one-micron

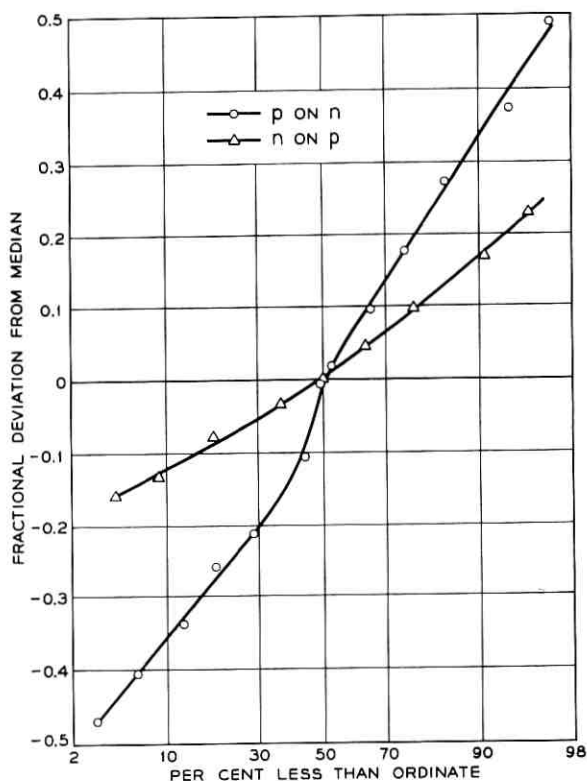


Fig. 2 — Statistical distribution of the degradation constant K .

thickness which represents the sum of the depletion layer and the sensitive region of the front layer. The solid curve was computed from (1) with $L_0 = 146 \mu$ and $K = 1.22 \times 10^{-8}$. Capacitance measurements indicate a depletion layer thickness of 0.5μ , leaving 0.5μ for the front layer, which appears reasonable.

It is instructive to plot the deviation of the degradation constants K for the individual cells as compared to the median value. Such a plot is shown in Fig. 2. It can be seen that the n-on-p cells show a tighter distribution than the p-on-n cells. This suggests that there may be an uncontrolled variable in the bulk n-type material which affects the

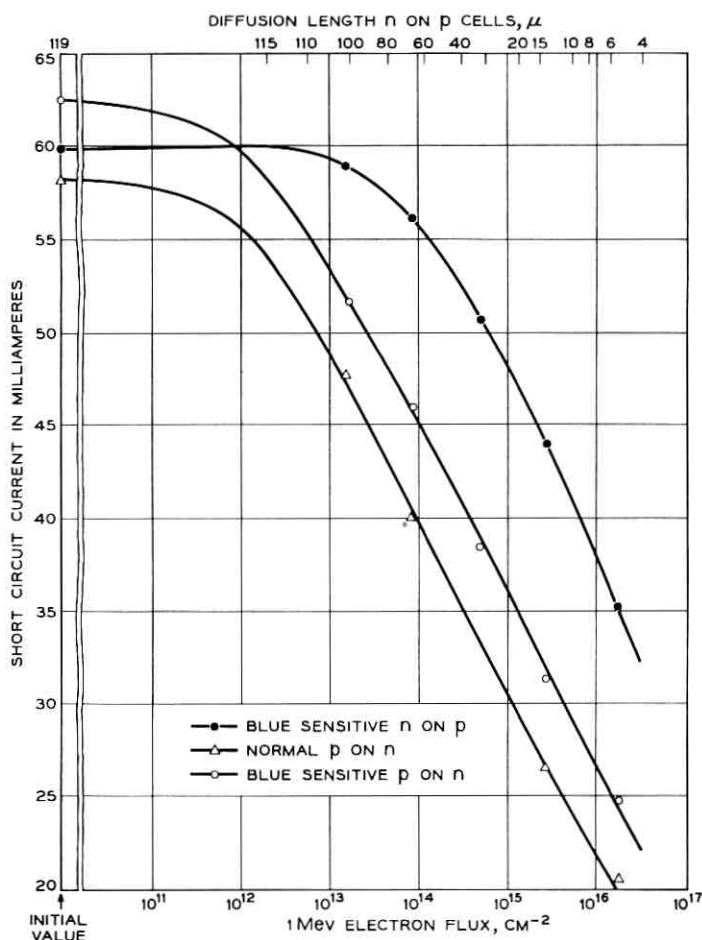


Fig. 3 — Outer space short circuit current versus 1-Mev electron flux.

sensitivity to particle radiation. This variable could possibly be the oxygen content of the material.

For the three types of solar cells the predicted mean outer space short-circuit current as a function of electron flux is given in Fig. 3. The results show that the radiation resistance of n-on-p cells is higher than that of p-on-n cells. The two types of cells reach the same short-circuit current for more than a factor of fifteen in electron exposure, which holds in spite of the fact that the n-on-p has initially a lower short-circuit current. This may be understood, qualitatively, by noting the greater rapidity with which the quantum efficiency, Figs. 4-6, degrades at the red end of the spectrum for the p-on-n cells as compared to the n-on-p. These curves also point out the significance of the spec-

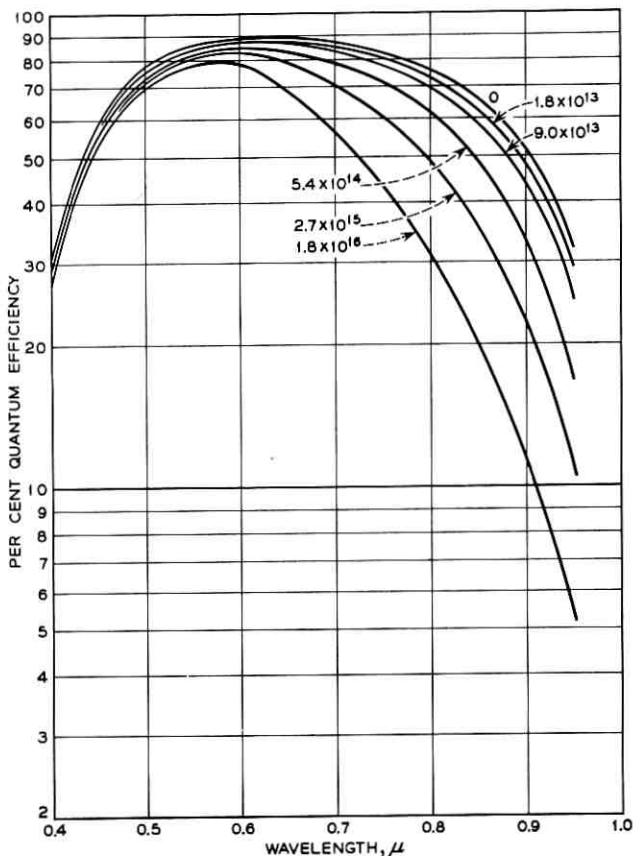


Fig. 4 — Per cent quantum efficiency versus wavelength after various levels of bombardment for blue-sensitive n-on-p cells.

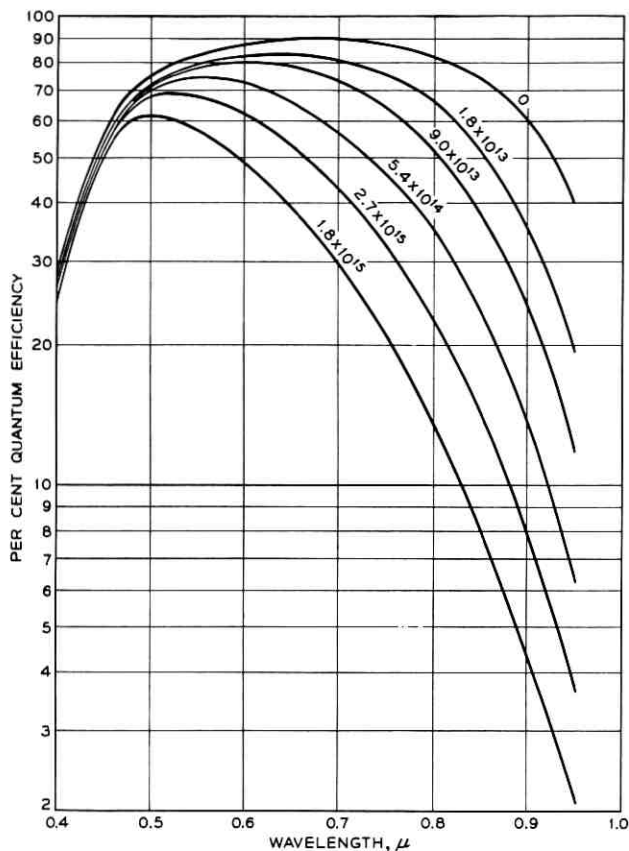


Fig. 5 — Per cent quantum efficiency versus wavelength after various levels of bombardment for normal p-on-n cells.

trum of the light source used to evaluate short circuit current degradation.

A better insight into the short-circuit current degradation may be obtained from the plots in Figs. 7-9. These show the measured contributions to the total short-circuit current for outer space sunlight falling into the indicated wavelength intervals as a function of bombardment flux. Since these contributions were measured by means of narrow-band filters centered in the indicated intervals one can compare the results with theoretical computations for monochromatic light.⁴ The solid curves are the results of such computations carried out by means of (2) normalized to the initial experimental short circuit current contribution.

$$I_i = I_{Ni} \left[(1 - e^{-\alpha_i x_F}) + \frac{e^{-\alpha_i x_J}}{1 + \frac{1}{\alpha_i} \sqrt{\frac{1}{L_0^2} + K\Phi}} \right] \quad (2)$$

I_i = short-circuit current contribution for the i th wavelength.

α_i = absorption coefficient for the i th wavelength.

x_F = the effective front layer thickness for the collection of carriers produced by penetrating radiation.

x_J = junction depth.

I_{Ni} = normalization factor for the i th wavelength.

In (2), the first term within the large brackets represents the contribution due to carriers collected from the total front layer, and the second

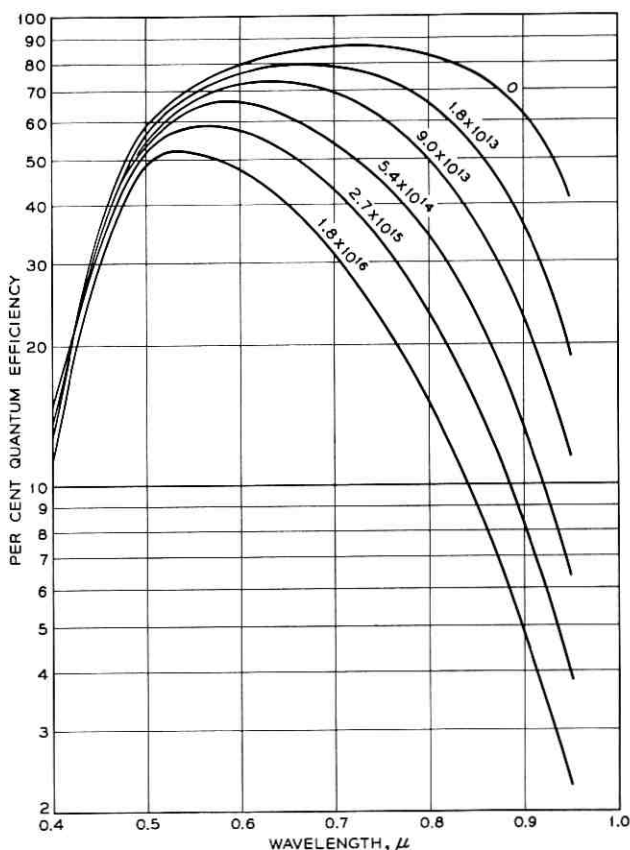


Fig. 6 — Per cent quantum efficiency versus wavelength after various levels of bombardment for blue-sensitive p-on-n cells.

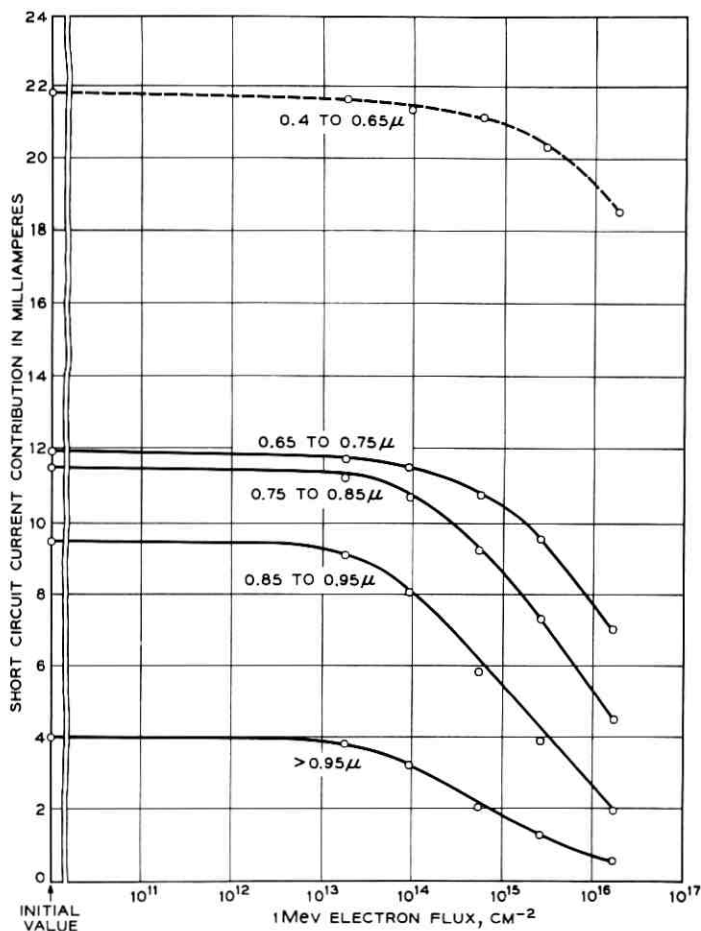


Fig. 7 — Contribution to outer space short-circuit current by sunlight in various wavelength intervals for n-on-p cells as a function of 1-Mev electron flux.

the contribution from the bulk. The effective thickness x_F is expected to be independent of wavelength only as long as $\alpha_i x_F \ll 1$. Thus, a comparison of the experimental points with (2) was made only for wavelengths greater than 0.65μ . This wavelength region is the one in which the greatest changes are produced by radiation. The agreement between the theoretical model and the experimental points is noted to be good.

The predicted maximum power for outer space sunlight as a function of flux is given in Fig. 10. The decrease in maximum power with bom-

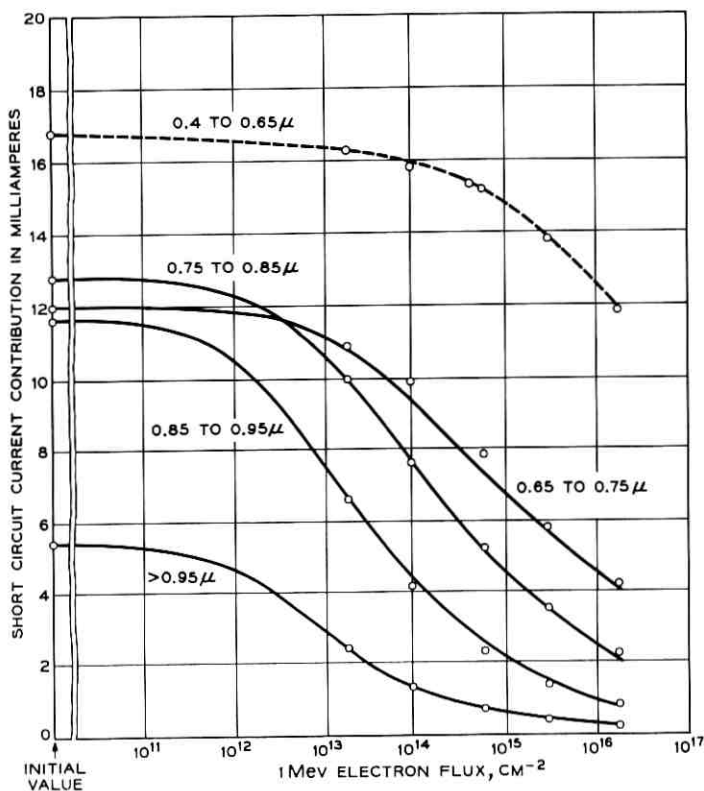


Fig. 8 — Contribution to outer space short-circuit current by sunlight in various wavelength intervals for normal p-on-n cells as a function of 1-Mev electron flux.

bardment is caused not only by the decrease in short-circuit current, but also by the degradation of the junction. A measure of the amount of this degradation is the decrease in the open-circuit voltage corresponding to a fixed short-circuit current. Fig. 11 is a plot of the open-circuit voltage corresponding to a short-circuit current of 50 ma as a function of 1-Mev electron flux.

These results may be analyzed by noting that over a limited range, the voltage-current relationship for the cell can be expressed by the formula

$$I = I_0 e^{(qV/nkT)} - I_{sc} \quad (3)$$

where I_{sc} is the short-circuit current, n is some number equal to or

greater than one, and I_0 is a saturation current. I_0 increases as the minority-carrier lifetime decreases, causing the maximum power to decrease even for a fixed short-circuit current. I_0 varies as τ^{-m} with $m = \frac{1}{2}$ or $m = 1$ depending on whether the saturation current is diffusion limited or due to space charge recombination, respectively.

At sufficiently high fluxes the inverse lifetime is proportional to the flux, so that the slopes in Fig. 11 yield the values $2.3 mnkT$. The slope for the n-on-p cells is 27 mv/decade and for the p-on-n 30 mv/decade. Both of these values are consistent with the pair of values $m \cong \frac{1}{2}$, $n \cong 1$. One can thus conclude that as far as the variation of diode

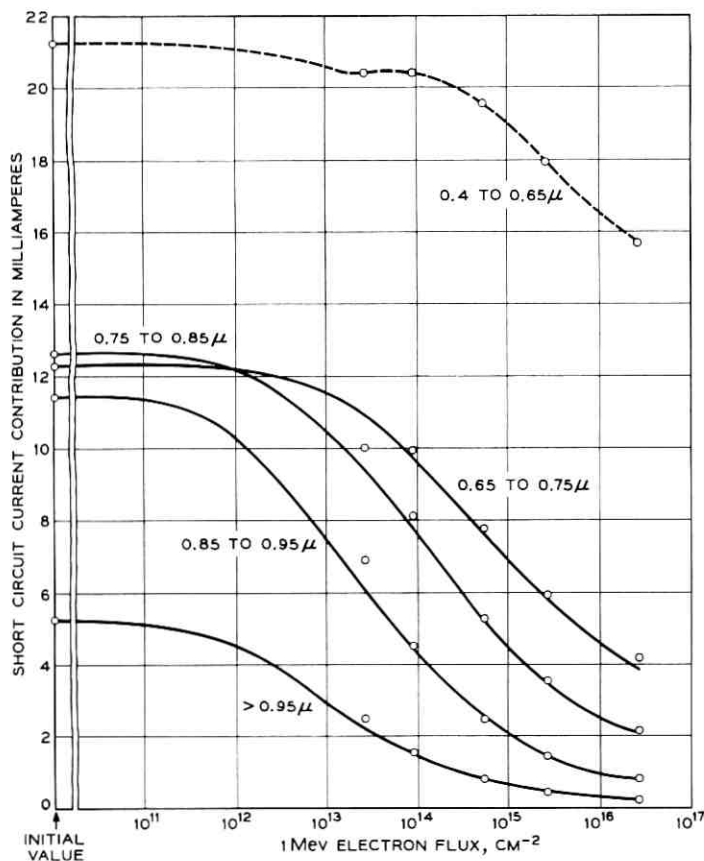


Fig. 9 — Contribution to outer space short-circuit current by sunlight in various wavelength intervals for blue-sensitive p-on-n cells as a function of 1-MeV electron flux.

characteristics with radiation damage is concerned, the solar cells behave as ideal diodes with diffusion-limited saturation currents (at current densities comparable to outer space short circuit currents).

Since the diffusion length variation with radiation follows a simple functional relation, its study is very suitable for finding relative radiation damage rates among different radiations. In essence, one needs to measure the K -values, as defined above, for protons as a function of

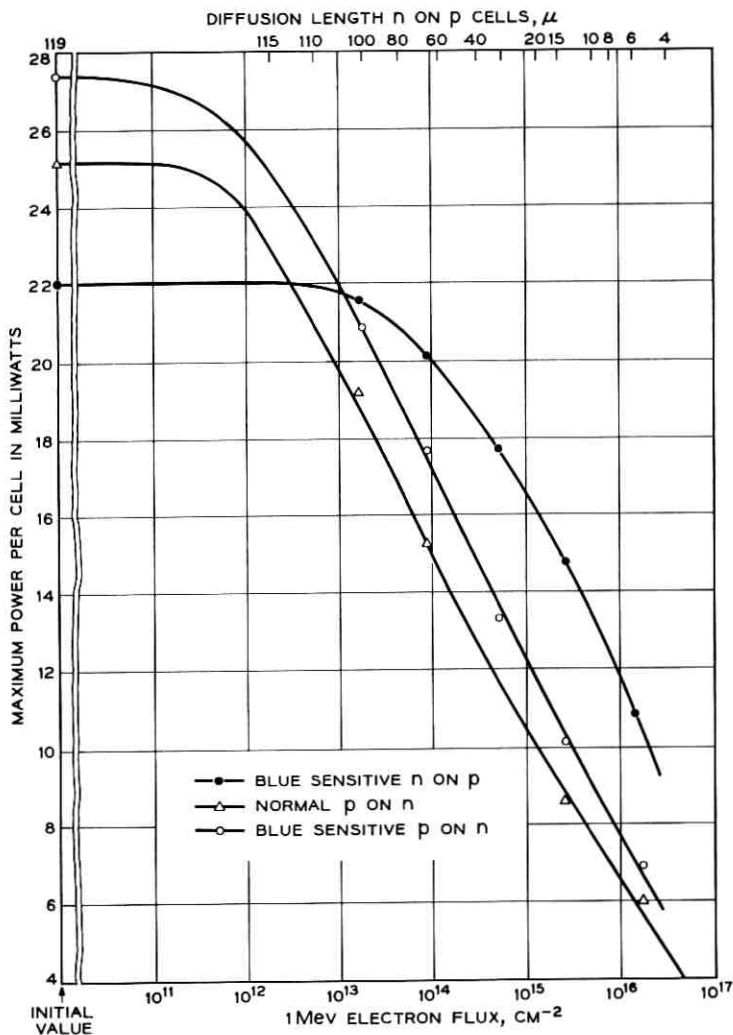


Fig. 10 — Maximum power versus 1-Mev electron flux.

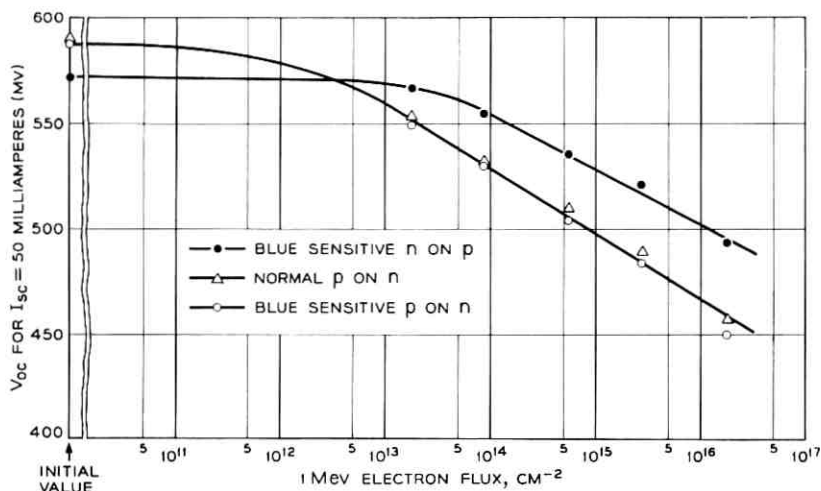


Fig. 11 — Open-circuit voltage corresponding to 50 ma short-circuit current versus 1-Mev electron flux.

energy, and for electrons as a function of energy and shielding thickness. The irradiation flux for each particle energy and shielding condition can then be converted to an equivalent 1-Mev electron flux by multiplying by the ratio of the K -values. This ratio can be termed as the “relative effectiveness” of the radiation with respect to the reference radiation, 1-Mev electrons in this instance. Thus, for example, if $R(E)$ is the ratio of K for n-on-p cells as a function of proton energy to K for 1-Mev electrons, and $\Phi(E)$ is the Van Allen Belt proton spectrum, averaged over some satellite orbit and shielding condition, in protons/(cm²sec Mev), then

$$\Phi_e = \int_0^{\infty} R(E)\Phi(E) dE \quad (3)$$

where Φ_e is the equivalent 1-Mev electron flux density in electrons/(cm² sec). All the figures contained in this report may then have their flux scales converted to time scales. It is to be noted, however, that such a conversion is not the same for n-on-p and p-on-n cells, because the ratio of proton to electron damage is different for n-type and p-type material.

V. CONCLUSION

The effect of radiation on the important parameters governing the performance of solar cells has been evaluated for three classes of cells

using 1-Mev electrons. The outer space short-circuit current, maximum power, junction characteristics, and spectral response have been presented quantitatively as a function of radiation flux along with the bulk minority-carrier diffusion length. The importance of the latter quantity is stressed as the unifying parameter which allows one to reduce fluxes of other types of radiation to equivalent fluxes of 1-Mev electrons.

In addition to supplying detailed engineering information, the results allow one to conclude the following:

The degradation in quantum efficiency at the longer wavelengths is quantitatively explained by the decrease in bulk minority-carrier diffusion length. The relatively small change in quantum efficiency at the short wavelengths suggests that surface recombination velocity and minority-carrier diffusion length in the front layer are not changed significantly by the radiation. The decrease in maximum power output is mainly due to the decrease in collection efficiency but to a smaller degree also to a degradation of the junction. The latter varies under forward bias with lifetime as an ideal diode with diffusion limited saturation current at current densities comparable to outer space conditions. The average relative fluxes required to achieve equal values for some of the parameters after prolonged bombardment, comparing only blue-sensitive cells, are shown in Table I.

TABLE I

Parameter	Flux Ratio (n-on-p relative to p-on-n)
Maximum power	9.5
Short circuit current	17
Diffusion length	72

VI. ACKNOWLEDGMENTS

The authors are very grateful to J. A. O'Sullivan, A. W. Norris, and K. R. Hill for their invaluable assistance in this work.

REFERENCES

1. Rosenzweig, W., *Rev. Sci. Inst.* **33**, 1962, p. 379.
2. Rosenzweig, W., *Diffusion Length Measurement by Means of Ionizing Radiation*, *B.S.T.J.*, **41**, Sept. 1962, p. 1573.
3. Gummel, H. K., and Smits, F. M., to be submitted for publication in the *B.S.T.J.*
4. Kleinman, D. A., *Considerations on the Solar Cell*, *B.S.T.J.*, **40**, January 1961, p. 85.

A Further Discussion of Stimulated Emission of Bremsstrahlung

By DIETRICH MARCUSE

(Manuscript received December 10, 1962)

In an earlier paper¹ we proved the existence of stimulated emission of bremsstrahlung for electrons moving in the vicinity of nuclei. However, no estimate of the available power was given at that time.

This paper extends the theory to the fourth order of perturbation theory, which allows one to estimate the available power from this process. We find that the available power increases proportional to the fifth power of the frequency, and that one might obtain power in the order of one watt at a frequency of 1000 gigacycles. The oscillation condition at these high frequencies is met by the passage of many slow electrons through a dense assembly of ions or atoms. Although these conditions are uncommon, stimulated emission of bremsstrahlung may play a role at microwave frequencies in very high current semiconductor experiments.

I. INTRODUCTION

In an earlier paper¹ we showed that stimulated emission of bremsstrahlung exists. This statement has the following meaning. Bremsstrahlung is the radiation which an electron emits by passing in the vicinity of a nucleus. We demonstrated that the emitted power into a specific mode of the radiation field is proportional to the energy density in that mode, which shows that we have indeed stimulated emission of radiation. It has to be expected that the emitted power can not be strictly proportional to the energy density, but that it must depend on the energy density in some nonlinear fashion. If this were not the case, the oscillation would not reach saturation and the energy density in the cavity of the oscillator would build up indefinitely. If we knew the nonlinear dependence of the emitted power on the energy density, we could predict the power output of a practical device.

Since it is very hard to find an exact solution of the problem, we will give an approximation by finding the next higher approximation of the perturbation theory, which will give us a term proportional to the square

of the energy density in addition to the already known linear term. This approximation will allow us to calculate the available power for electron currents which just barely exceed the current necessary to satisfy the oscillation condition. For higher electron currents this approximation will give not more than an order of magnitude estimate of the available power.

The result so obtained shows that the available power is proportional to f^5 (f = frequency) and is very low below 10 gigacycles (gc). However, the available power increases very rapidly, and a circuit designed to satisfy the oscillation condition at 1000 gc would deliver power in the order of one watt.

Since the use of a Coulomb potential appears to be an unnecessary restriction, we use a potential $V = e^{-\gamma r}/r$ which is an approximation to the potential of a neutral atom. The shielding effect of the electrons orbiting around the nucleus is taken into account by the factor $e^{-\gamma r}$. It turns out that stimulated emission will occur if $\hbar\gamma/mv \ll 1$ (with $2\pi\hbar = h$ = Planck's constant, m and v mass and velocity of the incident electron). This treatment neglects the interaction of the incident electron with the bound electrons in the atom. It is conceivable that additional emission or perhaps absorption of radiation may occur which is not included in this analysis. The bound electrons will certainly decelerate the free electron in flying through solid matter so that their presence poses additional serious problems.

Finally, in Section VII we discuss the advantages of using periodically distributed scattering centers. Periodic structures have the advantage that the emitted power becomes proportional to the square of the number of elements in the periodic arrangement. However, to utilize these at frequencies of 100 gc or higher requires the use of monoenergetic electrons, which are not easily available.

The utilization of stimulated emission of bremsstrahlung requires an arrangement which allows slow electrons to pass through dense assemblies of atoms or ions. This effect, most likely, plays a roll in semiconductor crystals to which high dc currents or current pulses are applied.

The following sections II to IV outline the fourth-order perturbation theory. Instead of the second quantized relativistic theory of the electron used by the authors of most textbooks, the problem is simplified by using a nonrelativistic approximation and first quantization only. The theory is presented with the purpose of showing the particular approximation used in deriving (34).

The reader who is not interested in the derivation of the theory may skip over sections II to IV and continue with section V. The expression

for the available power in section V is derived under the assumption that the incident electrons move parallel to the electric vector of the stimulating radiation field.

II. PERTURBATION THEORY

We will use quantum electrodynamics to derive the equation for the emitted power.

We simplify the problem by assuming that only one mode of the radiation field exists. The interaction of the electron with the photon vacuum will be neglected. This is justified as long as we are only interested in the stimulation effects and not in the spontaneous emission of bremsstrahlung. The state of the system including the electron and the radiation field will be described by a state function $\Phi(n, k)$; n designates the number of photons in the radiation mode while k refers to the electron propagation vector, which is related to the momentum $m\mathbf{v}$ of the electron by

$$\hbar\mathbf{k} = m\mathbf{v}. \quad (1)$$

The system is described by the Hamiltonian

$$H = H_0 + H_i. \quad (2)$$

H_0 is the Hamiltonian of the electromagnetic field plus the free electron. H_i is the interaction Hamiltonian, which is related to the interaction energy between the field and the electron.

According to Heitler²

$$H_i = H_1 + H_2 + V \quad (3a)$$

$$H_1 = -\frac{e}{mc} pA, \quad H_2 = \frac{e^2}{2mc^2} A^2, \quad V = -Ze^2 \frac{e^{-\gamma r}}{r}. \quad (3b)$$

The meanings of the symbols used are explained in the list of symbols at the end of this paper, Section VIII.

The vector potential is given by³

$$A = \frac{2\sqrt{\pi}c}{\sqrt{L^3}} (qe^{i\beta z} + q^*e^{-i\beta z}). \quad (4)$$

It is assumed that the x -direction coincides with the direction of the incident electron and the direction of the vector \mathbf{A} , which means that the electrons are incident parallel to the electric vector of the radiation field. The symbols q and q^* are the absorption and emission operators, respectively.

We assume that box normalization is being used. That means that the propagation constant k of the electron wave can assume only the values

$$k = \frac{2\pi\mu}{L} \quad (\mu = \text{integer}). \quad (5)$$

The electron wave function is given by

$$\psi = \frac{1}{\sqrt{L^3}} e^{i\mathbf{k}\cdot\mathbf{r}}. \quad (6)$$

Using the unperturbed wave functions of plane electron waves as zero-order approximation corresponds to the Born approximation. The state function Φ is the product of the photon state function and the electron wave function ψ . The time dependence of the state function is given by the Schroedinger equation.

$$i\hbar \frac{d\Phi}{dt} = (H_0 + H_i)\Phi. \quad (7)$$

It is more convenient to introduce interaction representation by making the transformation

$$\Phi' = \exp\left(\frac{i}{\hbar} H_0 t\right) \Phi \quad (8)$$

$$H_i' = \exp\left(\frac{i}{\hbar} H_0 t\right) H_i \exp\left(-\frac{i}{\hbar} H_0 t\right) \quad (9)$$

With these transformations (7) goes over into

$$i\hbar \frac{d\Phi'}{dt} = H_i' \Phi'. \quad (10)$$

The interaction energy is small compared to the energy of the noninteracting fields. We can therefore use an approximate iteration solution of (10). (We write henceforth H instead of H_i)

$$\Phi_P(t) = S\Phi_0(0) = (S_0 + S_1 + S_2 + \dots)\Phi_0(0) \quad (11)$$

where

$$S_0 = 1$$

$$S_1 = \frac{1}{i\hbar} \int_0^t d\tau_1 H'(\tau_1) \quad S_2 = \frac{1}{(i\hbar)^2} \int_0^t d\tau_1 \int_0^{\tau_1} d\tau_2 H'(\tau_1) H'(\tau_2)$$

$$S_3 = \frac{1}{(i\hbar)^3} \int_0^t d\tau_1 \int_0^{\tau_1} d\tau_2 \int_0^{\tau_2} d\tau_3 H'(\tau_1) H'(\tau_2) H'(\tau_3)$$

$$S_4 = \frac{1}{(i\hbar)^4} \int_0^t d\tau_1 \int_0^{\tau_1} d\tau_2 \int_0^{\tau_2} d\tau_3 \int_0^{\tau_3} d\tau_4 H'(\tau_1) H'(\tau_2) H'(\tau_3) H'(\tau_4).$$

The operator S is called the scattering matrix. The probability of finding a system, which at $t = 0$ is described by the wave function Φ_0 , after time t in the final state described by Φ_F , is given by the absolute square of the matrix element between these two state functions:

$$P = |(\Phi_F^*, S\Phi_0)|^2. \quad (12)$$

In order to evaluate the matrix element of the S -matrix it is necessary to convert the operator products $H'(\tau_1)H'(\tau_2)$ into matrix products. This can be done with the help of a complete set of state functions which are assumed to be eigenfunctions Φ_r of the unperturbed Hamiltonian H_0 with the eigenvalues

$$E_r = \frac{\hbar^2}{2m} k_r^2 + n_r \hbar \omega. \quad (13)$$

Making use of

$$(\Phi_r^*, H'(\tau)\Phi_s) = \exp\left[\frac{i}{\hbar}(E_r - E_s)\tau\right] (\Phi_r^*, H\Phi_s) \quad (14)$$

which follows from equation (9), we obtain, for example, for S_2

$$\begin{aligned} (\Phi_F^*, S_2\Phi_0) &= \frac{1}{(i\hbar)^2} \sum_r (\Phi_F^*, H\Phi_r)(\Phi_r^*, H\Phi_0) \int_0^t d\tau_1 \int_0^{\tau_1} d\tau_2 \\ &\cdot \exp\left[\frac{i}{\hbar}(E_F - E_r)\tau_1\right] \exp\left[\frac{i}{\hbar}(E_r - E_0)\tau_2\right]. \end{aligned} \quad (15)$$

Corresponding expansions hold for all the other S_i . The summation extends over all possible combinations of the products of all the free photon and free electron states.

III. EVALUATION OF THE MATRIX ELEMENTS OF THE S -MATRIX

It can easily be appreciated that for higher-order approximations the matrix elements of S_i become very complicated. Not only will an increasing number of nonvanishing terms appear in the summation but also the products of the individual matrix elements of H become increasingly involved since each H according to (3a) consists of three

terms, so that each term in the sum corresponding to S_i will contain 3^i terms.

Luckily, however, not all of these terms give a contribution. Let us, for the moment, assume that we are interested in the process of the emission of one photon. We see immediately that all terms containing only products of matrix elements of V give no contribution, since V does not change the photon number and since the scalar product of state functions with different photon number is zero. Also, all those products disappear which contain one term with the matrix element of H_2 and all others with V . H_2 changes the photon number either by 2 or 0, so that no state with one additional photon can result. These considerations, if applied to all possible products, reduce the number of those terms contributing to the matrix element S_i considerably.

In addition we will make one more approximation. Since we are interested in obtaining the nonlinear effects (nonlinear with respect to n), we do not want to improve the approximation of the result given in Ref. 1 as far as the potential V is concerned. In other words, we have obtained previously a result which was proportional to $e^6 n V$. By considering all terms in the higher-order approximations which contain one factor H_1 and an arbitrary number of factors containing V , we could obtain improvements in the previous calculation in the (symbolic) form

$$e^6 n (V + V^2 + V^3 + \dots).$$

However, these terms improving the previous approximation as far as the potential V is concerned will not give any information concerning nonlinear effects with respect to n . What we want instead is an approximation which results in

$$e^6 V n (1 + e^2 n + e^4 n^2 + \dots).$$

Therefore, we will consider only terms containing one factor V and neglect all those containing more than one.

Without going further into the details of the calculation, we will immediately quote the results of the calculation of the matrix elements which differ by one photon in the initial and final states. This fact will be indicated by a subscript 1 on the matrix elements. The matrix elements for the operators H_1 , H_2 and V are given by:

$$(\Phi^*(n \pm 1, \kappa), H_1 \Phi(n, k)) = -\frac{\sqrt{n}}{\sqrt{L^3}} \cdot \frac{e}{m} \sqrt{\frac{2\pi\hbar}{\omega}} \hbar k \delta_{\kappa_x, k} \delta_{\kappa_y, 0} \delta_{\kappa_z, \mp\beta} \quad (16)$$

$$(\Phi^*(n \pm 2, \kappa), H_2 \Phi(n, k)) = \frac{\pi}{L^3} \frac{e^2}{m} \frac{\hbar}{\omega} n \delta_{\kappa_x, k} \delta_{\kappa_y, 0} \delta_{\kappa_z, \mp 2\beta} \quad (17)$$

$$(\Phi^*(n,\kappa), H_2\Phi(n,k)) = 2 \frac{\pi}{L^3} \frac{e^2}{m} \frac{\hbar}{\omega} n \delta_{\kappa_x, k} \delta_{\kappa_y, 0} \delta_{\kappa_z, 0} \tag{18}$$

$$(\Phi^*(n,\kappa), V\Phi(n,k)) = -\frac{Ze^2}{L^3} \frac{4\pi}{|\mathbf{\kappa} - \mathbf{k}|^2 + \gamma^2}. \tag{19}$$

With these matrix elements we obtain

$$(\Phi_F^*, S_2\Phi_0)_1 = \mp \frac{4\pi}{(L^3)^{\frac{1}{2}}} \frac{Ze^3 n^{\frac{1}{2}}}{m\omega} \sqrt{\frac{2\pi\hbar}{\omega}} \frac{\kappa_x - k}{|\mathbf{\kappa} - \mathbf{k}|^2 + \gamma^2} \cdot \frac{\exp\left[\frac{i}{\hbar}(E_F - E_0)t\right] - 1}{E_F - E_0} \tag{20a}$$

$$(\Phi_F^*, S_3\Phi_0)_1 = -\frac{4\pi^2}{(L^3)^{\frac{1}{2}}} \frac{Ze^5 n^{\frac{1}{2}}}{m^2\omega^3\hbar} \sqrt{\frac{2\pi\hbar}{\omega}} \frac{1}{|\mathbf{\kappa} - \mathbf{k}|^2 + \gamma^2} \cdot \left\{ \mp 2i\omega t(\kappa_x - k) \frac{\exp\left[\frac{i}{\hbar}(E_F - E_0)t\right] - 1}{E_F - E_0} - 2i \frac{t}{\hbar} (\kappa_x + k - \kappa_x e^{\pm i\omega t} - k e^{\mp i\omega t}) \right\} \tag{20b}$$

$$(\Phi_F^*, S_4\Phi_0)_1 = \pm \frac{4\pi^2}{(L^3)^{\frac{1}{2}}} \frac{Ze^5 n^{\frac{1}{2}}}{m^3\omega^4} \sqrt{\frac{2\pi\hbar}{\omega}} \frac{\hbar}{|\mathbf{\kappa} - \mathbf{k}|^2 + \gamma^2} \cdot \left\{ (5k^2\kappa_x - 3k^3 - \kappa_x^2k - \kappa_x^3) \frac{\exp\frac{i}{\hbar}\left[(E_F - E_0)t\right] - 1}{E_F - E_0} + 2i \frac{t}{\hbar} [2\kappa_x^3 - 2k\kappa_x^2 + 2k\kappa_x(k - \kappa_x) \cos \omega t] \right\}. \tag{20c}$$

Of the two signs given in these equations, the upper sign refers to the case of emission and the lower sign of absorption of one photon. The value of $\kappa = |\mathbf{\kappa}|$ is also dependent on whether emission or absorption of one photon is being considered

$$\kappa = k \sqrt{1 \mp 2\epsilon} \tag{21a}$$

with

$$\epsilon = \frac{\hbar\omega}{mv^2}. \tag{21b}$$

The equations (21) follow from the conservation of energy $E_F = E_0$ with the help of (13).

The matrix elements (20b) and (20c) include the process of the virtual emission and absorption of one photon. While the final state Φ_F differs from the initial state Φ_0 by one photon, states with two additional photons appear in the intermediate steps of the calculation leading to (20b) and (20c). One may say, therefore, that these matrix elements correspond to a process either by which one photon is being absorbed and two photons are emitted, or by which two photons are emitted and one is absorbed. The final state has gained a photon in either case. In contrast to this, the matrix element (20a) corresponds to the simple emission or absorption of one photon without virtual emission and absorption processes taking place.

However, we have also to consider the case that two real photons are emitted or absorbed. This process leads to transition probabilities which are of the same order in e as the contributions which (20b) and (20c) will give to the emission or absorption probability of one photon.

The matrix elements corresponding to the emission or absorption of two real photons are given by:

$$(\Phi_F^*, S_2 \Phi_0)_2 = \pm \frac{2\pi^2 Z e^4 n \hbar \sigma_z}{L^6 m^2 \omega^2 c} \frac{1}{|\mathfrak{d} - \mathbf{k}|^2 + \gamma^2} \frac{\exp\left[\frac{i}{\hbar}(E_F - E_0)t\right] - 1}{E_F - E_0} \quad (22a)$$

$$(\Phi_F^*, S_3 \Phi_0)_2 = -\frac{4\pi^2 Z e^4 n \hbar}{L^6 m^2 \omega^3} \frac{(\sigma_x - k)^2}{|\mathfrak{d} - \mathbf{k}|^2 + \gamma^2} \frac{\exp\left[\frac{i}{\hbar}(E_F - E_0)t\right] - 1}{E_F - E_0} \quad (22b)$$

The subscript 2 on the matrix elements indicates that two photons have been emitted or absorbed. In (22) \mathfrak{d} is the propagation vector of the final electron with the energy

$$E_F = \frac{\hbar^2}{2m} \sigma^2 + (n \pm 2)\hbar\omega. \quad (23)$$

From $E_F = E_0$ follows

$$\sigma = k \sqrt{1 \mp 4\epsilon} \quad (24)$$

with ϵ of equation (21b).

IV. THE PHOTON CREATION RATE

We can now calculate the probabilities for the emission of one and also for the simultaneous emission of two photons. More important than the transition probability is the transition probability per unit time. According to Heitler⁴ this is obtained by summing the transition probability P over all possible values of the energy of the final state and dividing by the time t .

$$w = \frac{1}{t} \sum_F P = \frac{1}{t} \int_0^\infty P \rho_F dE_F. \quad (25)$$

ρ_F is the number of states per unit energy. We keep the energy of the radiation field fixed and allow the final energy of the electron to vary, disregarding conservation of energy for the moment. It turns out that conservation of energy is automatically assured since P contains a δ -function, as will be shown.

The number of states in the box of size L^3 (box normalization) turns out to be

$$\rho_F = \frac{mL^3 k_F}{\hbar^2 (2\pi)^3} d\Omega. \quad (26)$$

k_F is the magnitude of the electron propagation vector, which is either κ or σ depending on whether we consider the one- or two-photon process. $d\Omega$ is the element of solid angle into which the electron is scattered.

The probability P_1 for the emission of one photon is given by (12) and (20)

$$P_1 = |(\Phi_F^*, S_2 \Phi_0)_1 + (\Phi_F^*, S_3 \Phi_0)_1 + (\Phi_F^*, S_4 \Phi_0)_1|^2. \quad (27)$$

We neglect all terms of orders higher than e^8 . Two important terms will occur in (27). One is

$$\begin{aligned} t \frac{\exp \left[\frac{i}{\hbar} (E_F - E_0)t \right] - 1}{E_F - E_0} &= 2it \exp \left[\frac{i}{2\hbar} (E_F - E_0)t \right] \\ &\quad \cdot \frac{\sin (E_F - E_0) \frac{t}{2\hbar}}{E_F - E_0} \\ &= 2\pi i t \delta(E_F - E_0). \end{aligned} \quad (28)$$

The other one is

$$\begin{aligned}
 \left(\frac{\exp \left[\frac{i}{\hbar} (E_F - E_0)t \right] - 1}{E_F - E_0} \right)^2 &= 2\pi i \frac{\exp \left[\frac{i}{\hbar} (E_F - E_0)t \right] - 1}{E_F - E_0} \\
 &\cdot \exp \left[\frac{i}{2\hbar} (E_F - E_0)t \right] \quad (29) \\
 &\cdot \delta(E_F - E_0) \\
 &= 2\pi i \frac{t}{\hbar} \delta(E_F - E_0).
 \end{aligned}$$

The limiting process

$$\lim_{k \rightarrow \infty} \frac{\sin xk}{x} = \pi \delta(x)$$

can be taken since, even for relatively short times, we will have $E_0 \cdot \frac{t}{\hbar} \gg 1$. The somewhat daring calculation in (29) is suggested by Heitler⁴ and Mandl.⁵

The same calculations have to be performed with P_2 , the probability for the emission of two photons.

We see that P is proportional to t , so that $w = 1/t(\sum P)$ will be independent of time. P is also proportional to a δ -function which guarantees the conservation of energy. If we write

$$P = 2\pi \frac{t}{\hbar} |K_{F0}|^2 \delta(E_F - E_0)$$

we obtain

$$w = \frac{2\pi}{\hbar} |K_{F0}|^2 \rho_F. \quad (30)$$

This equation was taken as the starting point of the previous paper [Ref. 1, equation (1)].

We will not write down the explicit expressions for w_1 and w_2 but will go immediately to the photon creation rate $\Delta N'$. $\Delta N'$ denotes the number of emitted photons per second and, according to equation (15) of Ref. 1, is given by:

$$\Delta N' = (w_1 + 2w_2) \frac{L^3}{v} N_e N_n. \quad (31)$$

By substituting (30) and (27) and its equivalent for the two-photon process in (31), we obtain:

$$\begin{aligned}
\Delta N_{e'} = \frac{8\pi Z^2 e^6 N}{vm\omega^3 \hbar^2} d\Omega & \left\{ \frac{\kappa(\kappa_x - k)^2}{(|\boldsymbol{\kappa} - \mathbf{k}|^2 + \gamma^2)^2} \right. \\
& + \frac{2\pi e^2 N \hbar}{m^2 \omega^3} \left[\frac{\kappa(k - \kappa_x)[3(\kappa_x^3 - k^3) + 5k\kappa_x(k - \kappa_x)]}{(|\boldsymbol{\kappa} - \mathbf{k}|^2 + \gamma^2)^2} \right. \\
& \mp \frac{2m\omega}{\hbar} \frac{\kappa(\kappa_x^2 - k^2)}{(|\boldsymbol{\kappa} - \mathbf{k}|^2 + \gamma^2)^2} + \frac{1}{2} \frac{\sigma(\sigma_x - k)^4}{(|\boldsymbol{\delta} - \mathbf{k}|^2 + \gamma^2)^2} \\
& \left. \left. \mp \frac{1}{2} \frac{\beta\sigma\sigma_2(\sigma_x - k)^2}{(|\boldsymbol{\delta} - \mathbf{k}|^2 + \gamma^2)^2} \right] \right\} N_e N_n. \tag{32}
\end{aligned}$$

Terms proportional to $\cos \omega t$ have been neglected in (32) since, if the time average over one period of the oscillation is taken, no contribution from these terms would result.

$\Delta N_{e'}$ is the number of emitted photons if we ask only for the probability of photon emission and take the corresponding values for κ and σ and the upper signs. $\Delta N_a'$ is the number of absorbed photons. In order to obtain the net number of actually emitted photons, the difference $\Delta N' = \Delta N_{e'} - \Delta N_a'$ has to be taken.

$\Delta N_{e'}$ of equation (32) implies also that we are interested only in those photons which are emitted while the electron is scattered into a certain direction of space given by $\boldsymbol{\kappa}$ or $\boldsymbol{\delta}$ into the solid angle $d\Omega$. In order to get all the photons, we have to calculate

$$\Delta N = \int (\Delta N_{e'} - \Delta N_a') d\Omega. \tag{33}$$

It is interesting to note that the last two terms in (32), stemming from the two-photon process, give only a negligible contribution to (33) and can be neglected.

We obtain†

$$\frac{\Delta N}{N} = \frac{8Z^2 e^6 N_e N_n}{m^3 v^4 f^2} \ln \frac{2}{\sqrt{\epsilon^2 + \eta^2}} \left(1 - \frac{e^2 N v^2}{\pi \hbar f^3} \right) \tag{34}$$

with $\eta = \gamma/k \ll 1$ and $\epsilon = hf/mv^2 \ll 1$. Equation (34) holds as long as

$$\frac{e^2 N v^2}{\pi \hbar f^3} \ll 1. \tag{34a}$$

Equation (34) is identical to equation (32) of Ref. 1 if $\eta = 0$ in the limit $N \rightarrow 0$. [$\ln(2/\epsilon) - 1$ has been approximated by omitting the 1.] However, we now have obtained an expression for the photon creation

† All equations in this paper are written in electrostatic c.g.s. units.

rate $\Delta N/N$, which depends on the energy density hfN of the radiation field. For increasingly larger N , the creation rate $\Delta N/N$ becomes decreasingly smaller. We therefore have the possibility of computing the available power.

V. ESTIMATE OF THE AVAILABLE POWER

The power P_r radiated from a cavity is given by

$$P_r = 2\pi h f^2 N V \left(\frac{1}{Q_L} - \frac{1}{Q_i} \right). \quad (35)$$

The energy density hfN in the cavity will build up to the point where the number of created photons ΔN equals the number of photons which are radiated from the cavity and absorbed by its walls. We obtain for the oscillation condition

$$\frac{\Delta N}{N} = \frac{2\pi f V}{Q_L}. \quad (36)$$

Substituting (34) into (36), solving for N , and substituting into (35), we obtain

$$P_r = \frac{2\pi^2 h^2 f^5 V}{e^2 v^2} \left(\frac{1}{Q_L} - \frac{1}{Q_i} \right) \left[1 - \frac{\pi m^3 v^4 f^3 V}{4e^2 Z^2 Q_L N_n N_c \ln \frac{2}{\sqrt{\epsilon^2 + \eta^2}}} \right] \quad (37)$$

(c.g.s. units are being used).

Because of (34a), this equation holds if the expression in the brackets is much smaller than 1 (but larger than 0).

As soon as the product $N_n N_c$ is large enough to make $P_r \geq 0$, the cavity oscillates. The oscillation condition is, therefore

$$N_n N_c \geq \frac{\pi m^3 v^4 f^3 V}{4e^2 Z^2 Q_L \ln \frac{2}{\sqrt{\epsilon^2 + \eta^2}}}. \quad (37a)$$

It is understood that the power carried by the incident electron beam must be substantially larger than the power calculated from (37).

The most surprising fact about (37) is its dependence on the fifth power of the frequency. This means that at low frequencies the power obtainable from stimulated emission of bremsstrahlung is small.

Table I illustrates the situation for a practical example. We assume that we use bare nuclei with charge $Z = 1$. Correspondingly, we have to take $\eta = 0$. Furthermore, we assume that the expression in the bracket of (37) has the value of 0.1 and that $Q_L = 10^4$, $Q_i \gg Q_L$, $v = 2 \times 10^8$

TABLE I

f	P_r	$N_e N_n$	N_n/V
10 gc	10^{-10} w	10^{33} cm $^{-2}$ sec $^{-1}$	10^{15} cm $^{-3}$
100 gc	10^{-5} w	10^{37} cm $^{-2}$ sec $^{-1}$	10^{18} cm $^{-3}$
1000 gc	1 w	10^{40} cm $^{-2}$ sec $^{-1}$	10^{21} cm $^{-3}$

cm/sec corresponding to an accelerating potential of 10 volts. The cavity volume is taken as $V = 10$ cm 3 .

The values for $N_e N_n$ given in Table I are the products of nuclei and electron density current necessary to satisfy the oscillation condition (37a). N_n/V in the last column of Table I is the density of nuclei which results from $N_e N_n$ with $V = 10$ cm 3 if N_e is chosen to equal a current density of 0.1 amp/cm 2 .

We see that the available power is very low at $f = 10$ gc and reaches interesting values for $f = 1000$ gc. However, the required density of nuclei (or ions) also increases very rapidly with increasing frequency. Equation (37), which was used to compute the values for the available power in Table I, is an approximation. It does not hold for $N_e N_n \rightarrow \infty$. However, it seems reasonable to believe that the results of Table I are correct to the order of magnitude. We can not hope to obtain one milliwatt of power if our theory predicts 10^{-10} w.

VI. USE OF NEUTRAL ATOMS AS SCATTERING CENTERS

We derived (37) under the assumption that the electrons are scattered by a potential

$$V = -Ze^2 \frac{e^{-\gamma r}}{r}.$$

According to the Thomas-Fermi statistical model of the electron 6

$$\eta = \frac{\gamma}{k} = \frac{me^2 Z^{\frac{1}{2}}}{\hbar^2 k} = \frac{e^2 Z^{\frac{1}{2}}}{\hbar v}. \quad (38)$$

Equations (37a) and (38) allow us to calculate the required minimum number of atoms necessary to achieve self-sustained oscillations. It has to be remembered, however, that we have completely neglected the interaction of the incident electron with the bound electrons in the scattering atom. The only way the bound electrons enter the picture is by effectively shielding the Coulomb potential of the nucleus. In reality, the incident electron will excite the bound electrons into higher energy states and will also ionize some of the atoms. It might be that the process of ionization is accompanied by either stimulated emission or absorption

of radiation, just as the process of scattering is accompanied by a net stimulated emission of radiation if the incident electrons fly more or less parallel to the electric vector of the stimulating radiation field.¹

For all practical cases we will have

$$\epsilon \ll \eta$$

and therefore we obtain as the oscillation condition from (37a)

$$N_e N_n \geq \frac{\pi m^3 v^4 f^3 V}{4Z^2 e^6 Q_L \ln \left(\frac{2\hbar v}{e^2 Z^4} \right)}. \quad (39)$$

Table II lists the product $N_e N_n$ of (39), the atom density for an electron current of 100 ma/cm² and the velocity v as well as the corresponding accelerating potential U as a function of frequency for three different values of Z . The velocity is chosen so that $\eta = 0.2$ [equation (38)]. Z is equivalent to the order number of the atom in the periodic table of elements. We have again assumed $Q_L = 10^4$ and $V = 10$ cm³.

VII. DISCUSSION OF THE FEASIBILITY OF A BREMSSTRAHLUNG OSCILLATOR

The examples given in the last section show the problems involved in building an oscillator utilizing stimulated emission of bremsstrahlung. At low frequencies, where it is easy to satisfy the oscillation condition, the available power is very low. The available power reaches useful proportions in the region of 100 gc and becomes abundant at frequencies higher than 1000 gc. However, the required number of scattering nuclei or atoms becomes very high.

There is a possibility of using periodic arrangements of scattering centers rather than the randomly distributed atoms or nuclei considered here. The author has considered scattering by a string of nuclei (or atoms) which are arranged on a straight line parallel to both the direc-

TABLE II

f	Z	$N_e N_n$	N_n/V	v	U
10 gc 100 gc 1000 gc	1	3×10^{37} cm ⁻² sec ⁻¹ 3×10^{40} cm ⁻² sec ⁻¹ 3×10^{43} cm ⁻² sec ⁻¹	4.8×10^{19} cm ⁻³ 4.8×10^{22} cm ⁻³ 4.8×10^{25} cm ⁻³	1.1×10^9 sec	345 volts
10 gc 100 gc 1000 gc	10	6.5×10^{36} cm ⁻² sec ⁻¹ 6.5×10^{39} cm ⁻² sec ⁻¹ 6.5×10^{42} cm ⁻² sec ⁻¹	1×10^{18} cm ⁻³ 1×10^{21} cm ⁻³ 1×10^{24} cm ⁻³	2.37×10^9 sec	1600 volts
10 gc 100 gc 1000 gc	50	2.2×10^{36} cm ⁻² sec ⁻¹ 2.2×10^{39} cm ⁻² sec ⁻¹ 2.2×10^{42} cm ⁻² sec ⁻¹	3.5×10^{17} cm ⁻³ 3.5×10^{20} cm ⁻³ 3.5×10^{23} cm ⁻³	4.05×10^9 sec	4660 volts

tion of the incident electrons and the electric field vector of the stimulating radiation field. The result was that, if the distance between successive atoms is $d = v/f$, the number N_n of scattering centers entering (34), or (39) in case of a random distribution, is replaced by N_n^2 . Scattering by periodically arranged atoms thus greatly enhances the effect.

However, if the atoms are not aligned in one straight line but rather are in periodically arranged planes, N_n has to be replaced by nN_n , where N_n is the total number of atoms, while n is the number of planes. If 100 periodically arranged planes filled with scattering atoms were used, the number N_n quoted in Tables I and II would be reduced by a factor of 100, and the corresponding atoms would have to be distributed over these 100 planes. Using a frequency of 100 gc and atoms with $Z = 50$ would require a total number of 4×10^{19} atoms or 4×10^{17} atoms in each plane. If a plane of 10 cm² area is used, its thickness would have to be of the order of 1000 angstroms.

There is a limit to the number of planes which can be used in practice. This limit is set by the requirement that all electrons have to pass the distance between all these planes in the same time interval. The spread in electron velocities which can be tolerated decreases in inverse proportion to the number of planes used.

One could use electrons moving in conduction bands of solids rather than free electron beams and let them scatter from impurities in the crystal. Since the concentration of ionized impurities in semiconductors can be very high and since extremely high current densities can be applied, it appears that stimulated emission of bremsstrahlung should occur.

VIII. LIST OF SYMBOLS

$\mathbf{A} = (A_x, 0, 0)$	vector potential of the electromagnetic field
$c = 3 \times 10^{10}$ cm/sec	velocity of light
$\delta_{\kappa,k}$	Kronecker δ symbol equals 1 if $\kappa = k$ and is 0 otherwise
E	energy of the physical system composed of RF field and electron exclusive of the interaction energy
$e = 4.803 \times 10^{-10}$ dyn ^{-1/2} cm	electron charge (or base of natural Logarithm)
$\epsilon = hf/mv^2$	
$\eta = \gamma/k$	
f	frequency of the stimulating RF field
Φ	state function of the physical system
γ	shielding factor of the potential of the neutral atom

$h = 2\pi\hbar = 6.624 \times 10^{-27}$ erg·sec	Planck's constant
H_0	Hamiltonian of the noninteracting system
H_i	interaction Hamiltonian
$i = \sqrt{-1}$	or sometimes used as subscript
$k = k_x$	propagation constant of the incident electron
$\kappa = \sqrt{\kappa_x^2 + \kappa_y^2 + \kappa_z^2}$	magnitude of propagation constant of the scattered electron after the emission of one photon
L	length of the fictitious box used for box normalization
$m = 9.11 \times 10^{-28}$ gram	electron mass
n	number of photons in box of volume L^3
$N = n/L^3$	photon density
N_e	number of electrons per second per cm ²
N_n	total number of scattering nuclei or atoms
$\omega = 2\pi f$	
$p = mv = \hbar k$	momentum of the incident electron
$p = i\hbar(\partial/\partial x)$	momentum operator
q	absorption operator
q^*	emission operator
Q_L	loaded Q of the resonant cavity
Q_i	intrinsic Q of the resonant cavity
r	radius of polar coordinate system
ρ_F	number of electron states per unit energy interval
$\sigma = \sqrt{\sigma_x^2 + \sigma_y^2 + \sigma_z^2}$	magnitude of the electron propagation vector after the emission of two photons
S	scattering matrix
t	time
U	acceleration voltage of the electron
v	velocity of the incident electron
V	volume of the resonant cavity, also used to describe the shielded Coulomb potential
Z	number of elementary charges on scattering nucleus

REFERENCES

1. Marcuse, D., Stimulated Emission of Bremsstrahlung, B.S.T.J. **41**, September, 1962, pp. 1557-1571.
2. Heitler, W., *The Quantum Theory of Radiation*, Oxford, 3rd Edition, 1957, p. 125, equation (7).
3. Ref. 2, p. 143, equation (19).
4. Ref. 2, p. 139.
5. Mandl, F., *Introduction to Quantum Field Theory*, Interscience Publishers, Inc., 1959, p. 110.
6. Schiff, L. I., *Quantum Mechanics*. McGraw Hill Book Co., Inc., 1955, p. 170.

Engineering of T1 Carrier System Repeated Lines

By H. CRAVIS and T. V. CRATER

(Manuscript received December 10, 1962)

The repeated lines for the T1 carrier system (24 voice channels, PCM) are cable pairs equipped with transistorized regenerative amplifiers. The line signal is a train of 1,544,000 bipolar pulse positions per second. The line engineering methods, which tell how to select cable facilities and where to place repeaters, are based on the theory and measurements reported here. Using typical cable data, specific examples are given of the design limits imposed by the principal types of interference: crosstalk from other T1 systems and noise originating in central offices.

I. INTRODUCTION AND PROBLEM STATEMENT

1.1 System Description

The T1 carrier system transmits 24 voice channels over two cable pairs, one for each direction of transmission. The channels are time-division multiplexed and encoded by PCM, so that the line signal is a train of 1,544,000 bipolar pulse positions per second. Transistorized regenerative amplifiers, powered directly over the transmission pairs, detect the pulses and retransmit them along the line. The amplifiers are placed in manholes or on telephone poles at a nominal spacing of 6000 feet, for 22-gauge paper-insulated cable. This conforms to the common "H" loading spacing. General equipment features have been described for an experimental system,¹ and further aspects of the T1 system are described in another paper.²

In this paper we give the underlying theory for engineering the repeated lines. The emphasis is on the principles that affect the system layout. Although some specific objectives are developed, we have not tried to list all of the factors that may affect a specific layout.

1.2 Repeated Line Engineering

1.2.1 The Span as a Line Engineering Unit

A major departure in T1 line engineering, compared to previous short-haul carrier systems, is the design of routes in units called *spans*, as com-

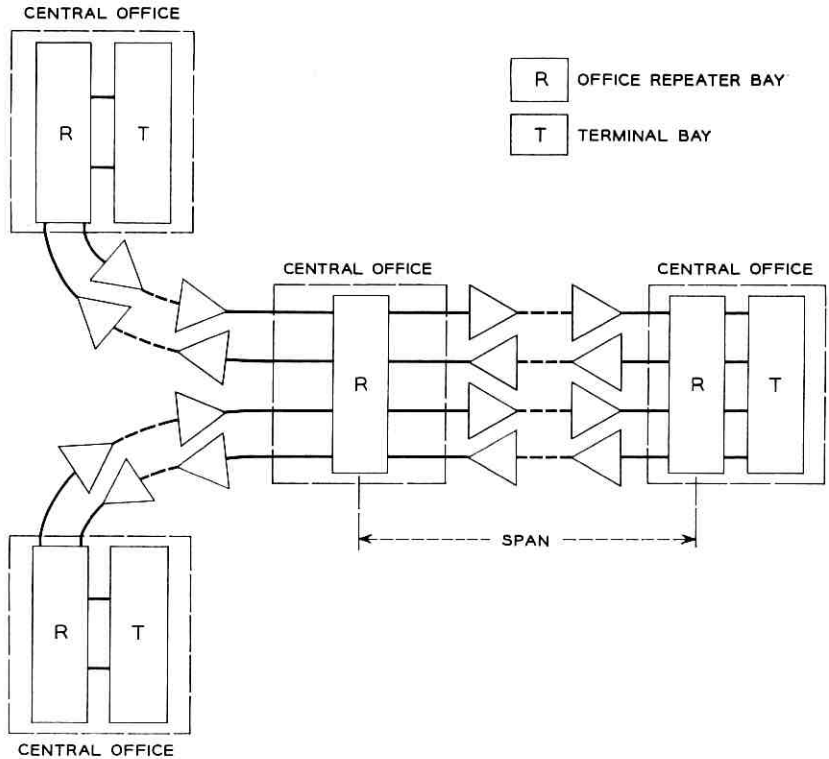


Fig. 1 — Spans.

pared to complete systems from terminal to terminal, or from one office served by the voice circuits to another. Referring to Fig. 1, a span is a collection of repeatered lines* between office repeater bays. A repeatered line for a particular system (terminal to terminal) is made up by interconnecting span lines at the offices along the route. It is simpler to administer, operate, and maintain the large numbers of telephone circuits that are needed in typical central office areas if the T1 lines are thought of as composing the spans to which they belong, rather than identifying the lines by their ultimate terminals. For example, any line in a span may serve as a spare for another, irrespective of the terminals served by the trunks using the spans. By building up the network on the span basis,

* Two regenerative amplifiers, with common power circuitry, make up a plug-in package called a repeater. Each repeatered line consists of two pairs of wires, one for each transmission direction, and the necessary amplifiers. Our "amplifier" is the same functional unit called a "repeater" by Mayo³ and a "reconstructive repeater" by Aaron.⁴

it is also easier to provide order wires for communication by maintenance personnel along the lines, to plan the fault-location requirements,² and to supply dc power for the repeaters. Our point of view, then, is to assure satisfactory pulse transmission in the lines of each span.

The prime markets for the T1 system are areas of high telephone development, such as the large networks in metropolitan areas. The majority of systems will provide trunks, e.g., local, tandem, and toll connecting trunks, between central office buildings in these areas. Typical cable lengths between buildings, corresponding to span lengths, are 2 to 10 miles. Terminal-to-terminal system lengths may be as great as 50 miles.

Message channels are needed in large numbers along T1 routes. Repeaters will be installed in a case suitable for location in manholes or on telephone poles. Each case will mount 25 repeaters, that is, 50 regenerative amplifiers as packaged. If the case is spliced into the main cable as in Fig. 2(a), 25 amplifiers serve as many pairs in one direction of transmission, and 25 the other; all 50 pairs are in the same cable, and this is called *one-cable operation*. In Fig. 2(b), all 50 amplifiers serve as many pairs for only one direction of transmission, and 50 more pairs in a second cable sheath, spliced to a second case, are needed for the opposite direction; this is *two-cable operation*.^{*} Each fully-equipped repeater case thus serves the lines in both directions of transmission for 600 message channels in one-cable operation, or in one direction for 1200 channels in two-cable operation. To allow for flexibility in pair assignment and to accommodate growth, pairs not immediately needed as T1 lines may be used for voice-frequency circuits by plugging into the repeater case special load coil cases or through connectors, which are interchangeable with the repeaters.

1.2.2 Objective

In T1 lines fidelity of transmission is measured by the error rate, which is the fraction of received pulse positions in which a pulse is present when none was transmitted, or vice versa. The error rate objective for a terminal-to-terminal line is based on two desiderata. First, the noise power in a speech channel caused by line errors must be negligible compared to noise and distortion from other sources. The performance of the terminal then is controlling.² Second, the distinctive, low-level noise pulses in a speech channel which result from line errors must occur infrequently enough to be unobjectionable to the user. The second is more

^{*} A repeater case is never spliced into more than one cable.

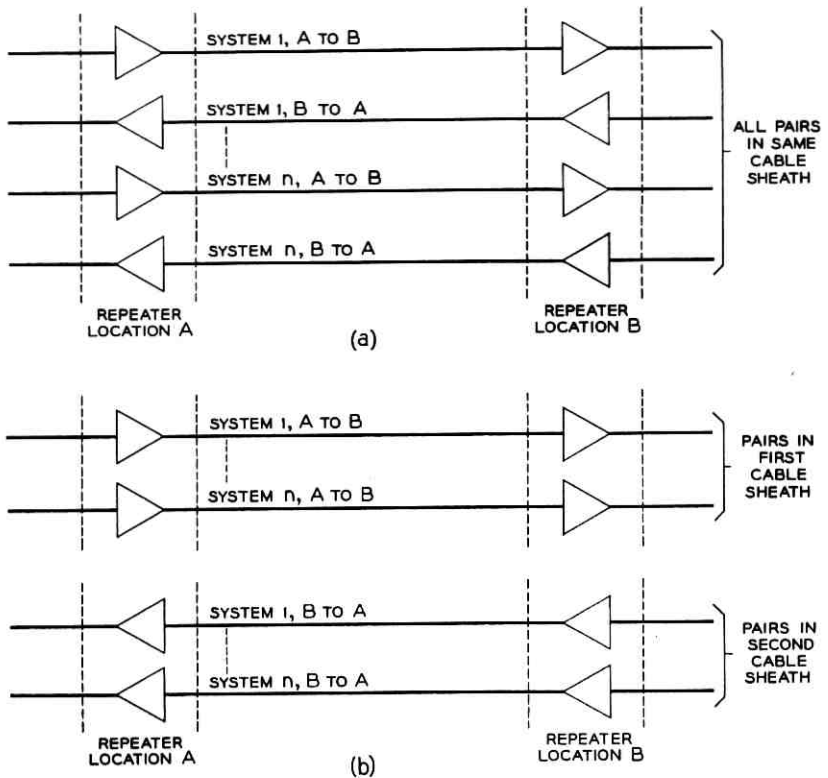


Fig. 2 — One- and two-cable operation.

difficult to satisfy than the first. Some tests conducted by W. L. Ross indicate that errors occurring at a rate of 10^{-6} are very difficult to detect by listening, and this rate has been made the system objective.¹

The goal of line engineering is satisfactory transmission over the lines of each span. The span objective is an error rate of 3×10^{-7} ; this allows three span lines in a system to have the maximum error rate. Since most span lines will have substantially lower error rates, the system objective will almost always be met, even in long systems. The system error rate of 10^{-6} is a conservative figure; an error rate of 10^{-5} will not seriously impair speech quality.

1.2.3 Further Requirements

Many aspects of the T1 system are novel; for example, the application of PCM. We wish to explore the theory that is relevant to line engineer-

ing, and then to reduce its results to manageable procedures for the layout of lines. Studies of the existing plant and trunk forecasts will show the numbers of voice channels needed between particular terminals and the available routes and cables. Starting with these, the procedures must tell how to lay out the T1 lines, or more specifically, how to choose the one- or two-cable mode (Fig. 2), how to choose the cable and pair assignments, and where to place the repeaters.

System layouts must come reasonably close to an economic optimum in the installed cost of repeaters. Measurements of cable properties, such as noise levels or crosstalk coupling losses, should not be necessary to proceed with the design. It should be possible to arrive at a workable layout without extensive calculation. In specific cases, the procedures may allow a number of alternatives, and the selection can be assumed to rest on the economics of the complete trunk plan.

1.3 *Types of Cables*

T1 carrier continues a trend in carrier telephone transmission toward the greater utilization, in terms of bandwidth, of cable originally installed in the outside plant for the transmission of voice frequencies. Prior examples are the N and ON systems.^{5,6} Each new system places its own demands on the cable plant. A natural development is that additions to this plant will be more and more influenced by potential carrier-frequency applications.

The present system is specifically intended for use on paired cables with paper strip, pulp or plastic conductor insulation. Load coils, building-out capacitors, and bridged taps must be disconnected from the pairs intended for carrier use. Cables of layer or unit construction with staggered pair twists will have satisfactory crosstalk properties, as will be seen below, but cables with nonstaggered twists are not acceptable. The preponderance of T1 applications will be on cables of larger sizes, say 200 pairs or more, although smaller cables will occasionally be used. A majority of installations, particularly in metropolitan centers, will utilize underground cable in ducts. In the outlying sections of cities, aerial cable is encountered; a field trial is presently under way to test the system performance in such cables, especially as regards the greater exposure to lightning and the wider temperature variations.

A considerable amount of 22-gauge cable is in place along the trunk routes involved; this gauge is therefore the design center for T1. Sections of gauges 19, 24, and 26 are also possible, and of mixed gauges where necessary.

More detailed cable characteristics that are desirable for T1 purposes are mentioned in the following sections.

II. AMPLIFIER PROPERTIES AFFECTING LINE ENGINEERING

2.1 Required Signal-to-Noise Ratio

An ideal binary pulse detection system^{7,8} can operate with noise amplitudes as great as half the received pulse amplitude. An error is made only when the noise value at the sampling instant is more than this amount and occurs with appropriate sign. A practical regenerative amplifier³ as built for the T1 system cannot operate with such a low signal-to-noise ratio. We assume that the probability of an error, that is, of retransmitting a pulse when none is received or vice versa, is the probability with which the noise amplitude exceeds one-fourth of the received pulse height. Production amplifiers are, of course, not all identical, and the required signal-to-noise ratio is taken as 12 db with a standard deviation of 1 db, based on observations of a number of samples.

2.2 Operating Signal Range

In discussing signal levels and line lengths, a convenient short cut is to replace the random pulse train, Fig. 3, with the periodic train of Fig 4.

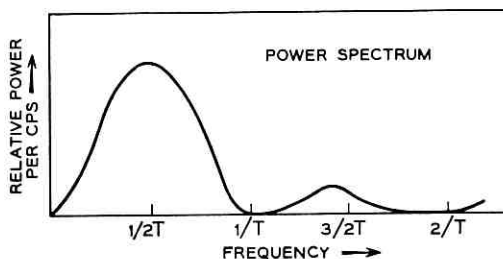
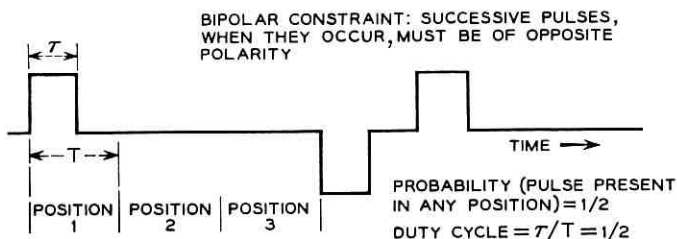


Fig. 3 — Random bipolar pulse train.

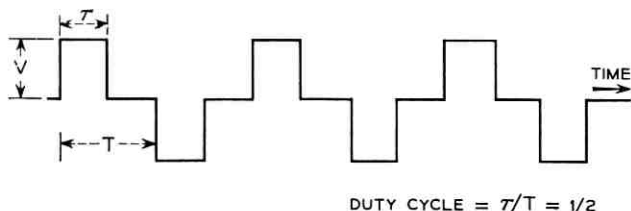


Fig. 4 — Continuous bipolar pulse train.

The random train, assuming balanced pulses, has no discrete frequency components.⁴ The maximum of its power spectrum is at about half the pulse repetition frequency, or 772 kc.* The continuous train has 81 per cent of its power at this frequency and the balance at its odd harmonics. The peak amplitude of the 772-kc component for the wave of Fig. 4 is $0.90 V$, or 2.7 volts for $V = 3$ volts at the amplifier output. For engineering purposes, the received pulse height may be found by calculating the peak amplitude of the received 772-kc component. Suppose, for example, that the cable pair between the output of one amplifier and the input of the next has 20 db attenuation at 772 kc. Then the pulse height at the sampling point, viz., at the equalized preamplifier output,³ is 0.27 volts, increased by the preamplifier gain at 772 kc. *We shall refer always to the attenuation at 772 kc and 55 F (degrees fahrenheit), when the term "loss" is used.*

The amplifier operates most effectively over a range of losses centered at 31 db, the loss of 6000 feet of 22-gauge paper insulated cable. In Fig. 5, the relative signal-to-noise ratio, at a constant error rate of about 10^{-6} , is plotted against the loss, for the particular case of near-end cross-talk interference. If we set a goal of 1 db departure from optimum, the figure shows that the loss may be 27 to 35 db. A wider range of line losses is made possible by building out shorter sections.³ Each amplifier has a buildout network at its input, selected on the basis of pair loss measurements at the time the repeater case is installed. The networks simulate the attenuation-vs-frequency characteristic of 22-gauge cable pairs, and are identified by their losses at 772 kc. For practical reasons, networks are supplied only in multiples of 2.4 db, as explained below.

The size of the loss steps for the networks is dictated by the loss range allowable for line and buildout together, 27 to 35 db, and by the variations due to temperature and to pair-to-pair differences. The increase in attenuation from 55 F to 100 F, the maximum temperature expected for

* The maximum is at 711 kc; the spectrum level at 772 kc is 98 per cent of this maximum.

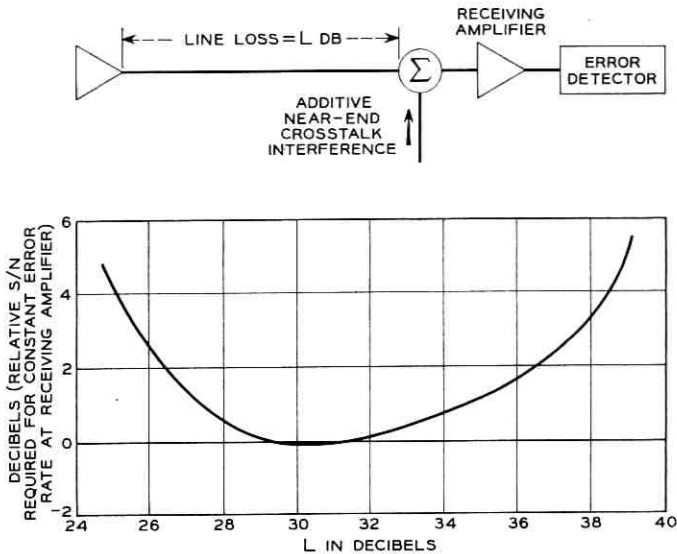


Fig. 5 — Effect of line loss on amplifier performance.

underground cable, is 1.3 db for 6300 feet* of 22-gauge cable. A comparable decrease in attenuation occurs when the temperature drops from 55 F to zero F, the minimum expected for underground cable. The standard deviation of pair losses is 0.6 db for this kind of section; therefore, the one per cent highest loss pair has an additional 1.5-db loss. So far we have accumulated a variation of ± 2.8 db. If the buildout networks are made in steps small enough to have the line plus network loss 31 ± 1.2 db, i.e., steps of 2.4 db, the desired range of 31 ± 4 db is achieved. Thus, networks of loss 0, 2.4, 4.8, . . . db are needed. In the experimental system,³ coarser steps of buildout loss were used. The list of networks for the commercial system has been filled in with the intermediate steps, to a maximum of 26.4 db. Line sections shorter than 32.2 db are built out to 31 ± 1.2 db. Generally all amplifiers in one direction in the same repeater section are equipped with the same buildout.

In making the preliminary layout, section losses are calculated from plant records showing the cable lengths and types. The average losses of common exchange cables are listed in Table I.^{9,10,11} For T1 calculations, we use the "engineering loss," which is the average loss increased by a small amount to account for month-to-month variations in cable produc-

* Maximum "H" loading spacing.

TABLE I — AVERAGE CABLE LOSSES, 772 KC, 55 F

Gauge	Insulation	Capacitance, microfarads/mile	Loss, db/mile
19	polyethylene	0.066	11.8
19	polyethylene	0.083	15.1
19	paper strip	0.066	15.2
19	paper strip	0.084	19.1
22	polyethylene	0.083	21.9
22	paper pulp	0.082	26.5
24	paper strip	0.072	29.5
26	paper pulp	0.079	39.2

tion. For 22-gauge cable with average loss of 26.5 db per mile, as an example, the engineering loss is 27.0 db per mile, or 32.2 db per 6300 feet. When a section is composed of mixed gauges of cable, or of high- and low-capacitance types of the same gauge, a "junction loss" is added to the engineering loss. The junction loss is based on the reflection loss at the junction, suitably reduced if only a short section of different gauge is present. It is not more than 0.3 db for common cable types.

2.3 Minimum Section Loss

The amplifier output and input impedances are not perfectly matched to the characteristic impedance of the cable pairs over the band of frequencies needed for pulse transmission. As a result, reflections of pulses may travel back from an amplifier input to the output circuit of the previous amplifier and interfere with the regeneration of pulses there, increasing the error rate. Experience has shown that this is serious only for amplifiers separated by less than 9 db of cable loss, a situation that may arise adjacent to central offices (cf. Section 4.5). For this reason, central office amplifiers are equipped with fixed pads of 100 ohms impedance and 3 db loss, at both input and output. These pads also reduce reflections from cable plugs and terminating cable discontinuities at the cable vault. Further, a minimum line loss of 6 db is required between the central office main distributing frame and the nearest outside repeater point. Sections not adjacent to central offices must have a minimum line loss of 9 db; in practice, these sections are normally long enough to meet this limit without special consideration. Operation with these minimum section losses* also helps in fault locating,² during which the amplifier is especially sensitive to interference from reflections.

* The two highest buildouts (Section 2.2) of 24 and 26.4 db were originally provided to complete the series. With a minimum line loss of 9 db, these buildouts are not used.

III. SOURCES OF INTERFERENCE

3.1 *General Remarks*

In the experimental system,⁴ a choice of fundamental importance was made: the transmission of bipolar pulses (see Fig. 3) with timing wave extraction at half the pulse repetition rate. A simplified model of crosstalk was assumed in Ref. 4 to estimate the relative effects of crosstalk on the timing wave that is derived at a disturbed amplifier, for various pulse transmission schemes. Experience with the experimental and prototype systems has confirmed the value of the earlier choice. In this paper, we deal with the effects of interference on the pulse detection process. In practical installations, errors due to additive interference at the decision point of the amplifier, rather than those due to timing wave jitter, limit the accuracy of transmission of the pulse train.

We have stated the performance criterion in Section 1.2.2 in terms of amplifier error rate. For ideal pulse detection, this rate may be found if we know the amplitude of the received pulses and the amplitude distribution of the interfering voltage at the point where the regeneration decision is made. As an experimental observation, the T1 amplifier operates closely enough to this ideal amplitude-sensitive fashion that we may base error rate calculations on this model (cf. Section 2.1). In this section the major sources of interference are identified, and in Section IV the limits placed on repeatered line layout by the controlling sources are derived.

Internal sources of noise are principally thermal noise and disturbances introduced by the power supply. It is apparent that voltages from these sources are too small to enter into line engineering.⁴ For example, the rms thermal noise voltage¹² in a band from zero to two megacycles per second, at a temperature of 300 degrees Kelvin and resistance 100 ohms (the characteristic impedance of 22-gauge cable at 772 kc) is 2 microvolts. The signal level (see Section 2.2), or received pulse height as equalized, at the end of a line section of 32.2 db loss is 0.066 volts, more than 90 db above the rms noise.

External sources induce voltages directly into the repeatered lines and indirectly via crosstalk coupling paths.* For example, central office noise affects T1 lines principally via crosstalk paths from noisy non-T1 pairs. Also, atmospheric static, lightning and radio transmitters induce voltages in exposed parts of the wire plant such as open wire extensions or sub-

* The term "crosstalk" refers to voltages induced in T1 pairs from nearby conductors. The resulting interference is not in general audible speech or even "babble," as in voice-frequency usage.

scriber drops, which, in turn, may join pairs in T1 cables. A further interference problem is the coordination of T1 with other transmission systems on cable pairs or other conductors in the same sheath. The possibility of *mutual* interference exists between T1 and other systems with similar frequency bands, particularly type N carrier (48 to 256 kc), type L carrier (308 to 8320 kc) and type A2A television (baseband to 4500 kc).

3.2 *Controlling Types of Interference*

To proceed in an orderly fashion, we must know the relative importance of the several sources of interference and then develop line engineering procedures which assure us of satisfactory operation under the most adverse amounts to be expected, of the controlling types of noise, in any combination. At this point, we draw heavily on experience in field trials of the experimental and prototype T1 systems, the former between Summit and South Orange, New Jersey,¹ the latter between Newark and Passaic, New Jersey, as well as on data gathered at Bell Telephone Laboratories and in operating telephone company areas. The conclusion, in brief, is that *crosstalk from other T1 systems* in the same cable and *central office noise* are the forms of interference that require specific engineering measures to combat.

Other sources do not seem at present to cause serious levels of interference. If future applications warrant it, these other sources may be taken into account in line engineering by methods similar to those given here for office noise and intersystem crosstalk. The effects of atmospheric electricity will be better known quantitatively after the aerial cable trial, mentioned in Section 1.3. To date, no instance of interference from radio transmitters has occurred. With regard to other transmission systems, coordination with type N carrier is excluded in existing plans for T1, primarily to avoid crosstalk to type N. Studies are now being made of T1 interference to or from type L carrier or type A2A video systems in the same cable sheath. The latter two systems operate on coaxials and shielded pairs, respectively, in the cores of some paired cables.

3.3 *Crosstalk from Other T1 Systems*

Near-end, far-end, and interaction crosstalk must be considered. Preliminary measurements³ by the Outside Plant Systems Engineering Department at Bell Telephone Laboratories showed that interaction crosstalk coupling losses are extremely high in multipair cables. For this reason, this kind of crosstalk has not been dealt with further.

3.3.1 Near-End Crosstalk

Near-end crosstalk was recognized as a significant problem in one-cable systems during the development and testing of the experimental system.⁴ Referring to Fig. 2(a), the high-level outputs of one set of amplifiers are coupled to the inputs of another set via near-end crosstalk paths, the most damaging of which are in the cable near the repeater case. It was primarily to lessen the effects of near-end crosstalk on amplifier timing that the bipolar pulse train of Fig. 3 was introduced. Its power spectrum peaks at about 772 kc (see Section 2.2), whereas a unipolar pulse train contains relatively more energy at higher frequencies, as well as certain discrete components. A secondary advantage is that the absence of a dc component in the bipolar case eases low-frequency transmission requirements.

The nature of near-end crosstalk interference has been studied with the laboratory setup^{14,15} of Fig. 6, in which crosstalk from random bipolar word generators is added to the input of a test amplifier to determine its effect on error rate. The disturbing word generators have independent random pulse train outputs; a typical one is shown in ideal form in Fig. 3. Each generator is timed by its own crystal oscillator. The signal word generator of Fig. 6 can be set to give a fixed repetitive pattern or a pseudo-random pattern in which every eighth position and the position following it are, respectively, a forced mark and a forced space. Patterns

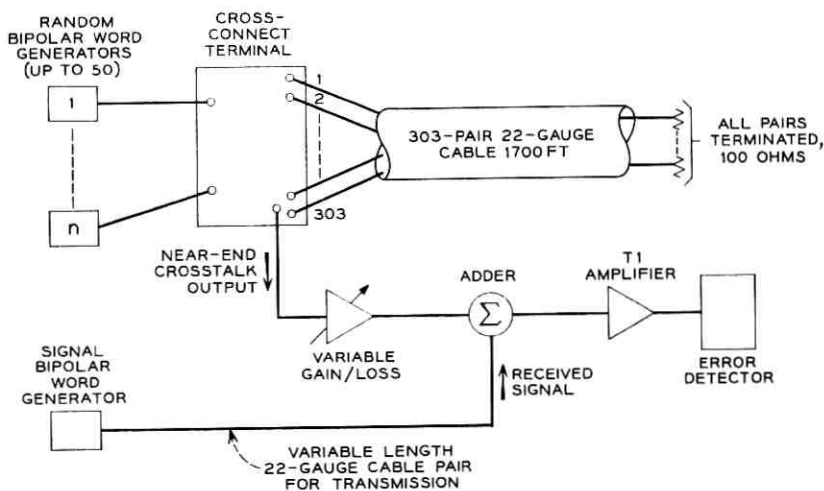


Fig. 6 — Laboratory setup for near-end crosstalk tests.

with very long runs of spaces must be avoided, as they may lead to self-oscillation in the receiving amplifier.³ In the present paper, our conclusions are not significantly affected by the choice of pattern for the received signal.

In Appendix A, a theory is developed for predicting the interfering power that arises when n systems operate in a cable. A knowledge of the mean value of this power, and of its standard deviation as we select combinations of n pairs, allows us to develop line engineering procedures, as is done in Section IV. As further pointed out in Appendix A, the laboratory tests have shown that the near-end crosstalk interfering voltage has a Gaussian amplitude distribution, out to several standard deviations.¹⁵ This is the reason we may calculate error rates in Section IV using only the rms value of the interference.

It is largely a result of experience with frequency-division carrier systems that many measurements of crosstalk losses in multipair cables have been made at selected frequencies in the band of interest. The near-end crosstalk coupling loss between typical cable pairs varies irregularly with frequency, as illustrated in Fig. 7. In fact, measurements across the frequency band on one pair combination are almost valueless; the statistical properties must be found by measuring a large number of pair combinations.⁹ In Table II, we have summarized data available on coupling losses at 772 kc in cables with staggered pair twists. Losses are given for pairs in 100-pair splicing groups, to correspond to pair selections in actual cables planned for T1 use. Using these data with the theory of Appendix A, as explained in Section 4.2, line engineering tables have been compiled to aid in the layout of one-cable systems.

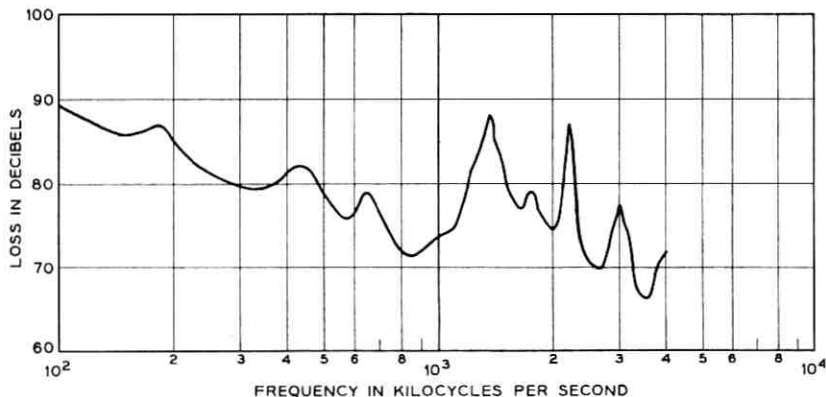


Fig. 7 — Near-end crosstalk loss vs frequency.

TABLE II — NEAR-END CROSSTALK COUPLING LOSSES AT 772 KC
FOR 22-GAUGE CABLES WITH STAGGERED PAIR TWISTS,
100-PAIR SPLICING GROUPS

Cable Construction	Mean, db	Standard Deviation, db
Unit		
Pairs in same group	82	11
Pairs in adjacent groups	90	9
Pairs in alternate groups	103	7
Layer		
Pairs in same group	75	9
Pairs in adjacent groups	83	8

These data apply, without correction for length, to sections at least 1000 feet long. The data were taken on sections about 6000 feet long.

3.3.2 Far-End Crosstalk

Far-end crosstalk coupling paths are always present among cable pairs in the same transmission direction. Three situations, illustrated in Fig. 8, are critical in T1 engineering. In Fig. 8(a), there are lines for the same direction of transmission in which the amplifiers are slightly misaligned. This may happen in one- or two-cable operation when repeater cases are spliced into the main cable at different points along the main cable. In Fig. 8(b), a greater misalignment is shown. This can occur when systems are added along a two-cable route, and there is not enough space at existing repeater locations to accommodate the additional repeater cases. It may be necessary, for example, to use the next manhole along the line for an underground installation. In parts (a) and (b) of the figure, the most damaging far-end crosstalk exposures, shown by arrows, are in the cable between an amplifier output and the nearest amplifier input. In Fig. 8(c), we have a junction, where lines from different systems enter the same cable; the figure shows only one direction of transmission. If L_A and L_B are unequal, say $L_B > L_A$, signals from the "A" amplifier output interfere more seriously with the "B" system than vice versa, by way of the coupling path shown.

The theory of far-end crosstalk is developed in Appendix B. As in the near-end case, we find the statistical properties of the interfering power when n systems are present in a cable. The assumption that the crosstalk voltage has a Gaussian amplitude distribution again allows us to estimate error rates based on interference power. The theory is applied to the layout of two-cable systems in Section 4.3, and to junctions in Section 4.4.

3.4 Central Office Noise

In a telephone switching office, a variety of devices and circuits create transient and repetitive currents with energy spread over a wide frequency band. The more common sources are relays and switches that interrupt direct or alternating currents, rectifier power supplies and ac power wiring, and lines carrying ringing signals and other plant and test tones. Carrier wiring inside the office, e.g., from main distributing frame to carrier equipment bays, can be isolated from these disturbances by shielding and by physical separation. Office noise can also find its way into carrier lines by secondary induction. That is, voice-frequency pairs within the office are exposed to office noise, and a number of these pairs may enter cables which contain carrier pairs in the outside plant. The noise voltages are coupled from the voice-frequency pairs into the carrier pairs by crosstalk, chiefly near-end and far-end.

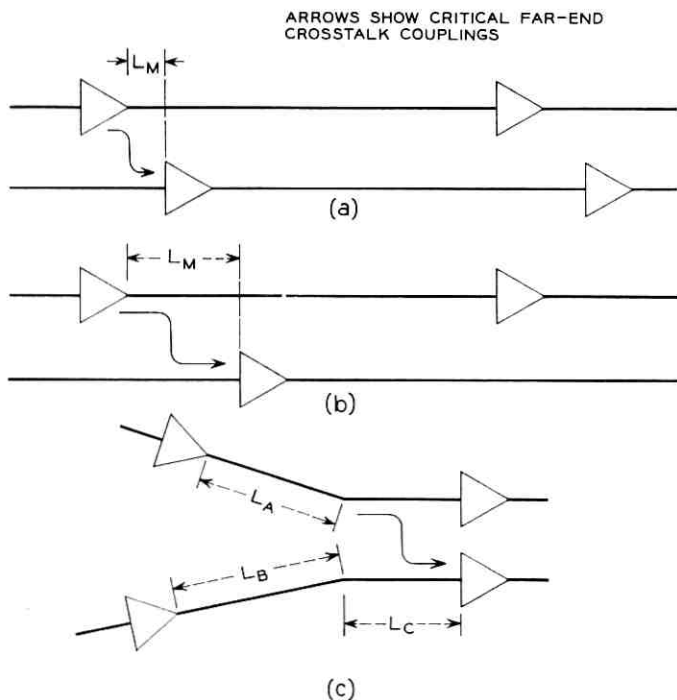


Fig. 8 — Far-end crosstalk situations: (a) sections with repeaters almost aligned, L_M small; (b) sections with greater misalignment of repeaters, L_M large; (c) junction.

Such secondary couplings are important in T1 lines only in the vicinity of central offices, as the noise on the noncarrier pairs, which is effective in the T1 band, is rapidly attenuated further out. In Appendix C some of the properties of office noise are described, and in Section IV a simple application of these facts is made to arrive at line engineering methods.

IV. DETAILED ENGINEERING OF REPEATER SECTIONS

4.1 *Error Rate Objectives for Repeater Sections*

In engineering a repeater section one has to take account of errors due to crosstalk and office noise so as to produce a sufficiently small contribution to the span error rate objective (Section 1.2.2) of 3×10^{-7} , while achieving nearly maximum economy in the use of repeaters.

The error rate caused by a combination of crosstalk and office noise is very nearly the sum of the error rates caused by the two kinds of interference separately (Appendix C). It is therefore possible to set an error rate objective for noise and one for crosstalk, and to find the separate design limits (cf. Section 4.5) that will meet these objectives. We have set the objective for error rate due to office noise in one section at 10^{-7} ; assuming two sections in a span to be fully exposed to noise, this assigns two-thirds of the span error rate. The remainder is assigned to crosstalk. Because of the extreme sensitivity of error rate to small differences in crosstalk loss or section loss, error rates due to crosstalk will vary radically from one section to the next. The section objective is an error rate of 10^{-7} for the worst one per cent amplifier. This means that in a five-section span about 95 per cent of the lines will meet the span objective. This is a reasonable performance level. Marginal lines may be set aside for use as spares, and the investment in still poorer lines may be salvaged by reusing the repeaters elsewhere and returning the cable pairs to voice-frequency use.

4.2 *One-Cable Systems*

4.2.1 *Design Section Loss*

We have mentioned in Section 3.3.1 that one-cable installations are subject to near-end crosstalk effects. We must limit the number of systems installed in a given cable, or the repeater section loss, or both, to stay within the error rate objective of Section 4.1. The quantitative limits are obtained as follows.

Suppose a repeater section has cable loss L db. Then the received pulse height at the preamplifier output is equivalent to $S + G - L$ dbm,* where S is the peak value of the 772 kc signal component at the transmitting amplifier output (cf. Section 2.2) in dbm, and G is the preamplifier gain at 772 kc. Let P_ϵ be the instantaneous noise power in dbm which is exceeded with probability ϵ . Then, for a required signal-to-noise ratio of A db (cf. Section 2.1), the error rate will be ϵ when $S + G - L - A - P_\epsilon = 0$. For near-end crosstalk, the interfering voltage has a Gaussian amplitude distribution (cf. Section 3.3.1). If its rms value for n interfering systems is equivalent to P_n dbm, then

$$P_\epsilon = P_n + 20 \log y(\epsilon) \text{ dbm} \quad (1)$$

where $y(\epsilon)$ is the (positive) quantity such that

$$\epsilon = (2\pi)^{-1/2} \int_y^\infty \exp(-t^2/2) dt. \quad (2)$$

Actually, S , L , A , and P_ϵ are statistical quantities. The mean value of S is $\bar{S} = 10 \log (2.7)^2/100 + 30 = 18.6$ dbm, and its standard deviation is $\sigma_s = 0.5$ db, owing to variations in the amplifier output circuits. The mean pair loss, \bar{L} , is a design parameter to be found here; we use a standard deviation $\sigma_L = 0.6$ db for \bar{L} about 30 db. The mean amplifier signal-to-noise ratio \bar{A} , as mentioned in Section 2.1, is 12 db, with $\sigma_A = 1.0$ db. As to P_ϵ , its mean value, introducing the mean \bar{P}_n as in (34), is

$$\begin{aligned} \bar{P}_\epsilon &= \bar{P}_n + 20 \log y(\epsilon) \\ &= 48 - m + \sigma + 10 \log (n/25) + 20 \log y(\epsilon) \end{aligned} \quad (3)$$

where m and σ are the mean and standard deviation, respectively, of the near-end crosstalk coupling losses at 772 kc, in db. The standard deviation of P_ϵ , which is the same as that of P_n , is underestimated by the theory of Appendix A. A value of $\sigma_{P_n} = \sigma_{P_\epsilon} = 3.2$ db has been determined experimentally.¹⁵

Accordingly, we now define another statistical quantity, the *margin at error rate* ϵ ,

$$M_\epsilon = S + G - L - A - P_\epsilon. \quad (4)$$

Substituting in (4) the values given, with $G = 23.7$ db, and $y = 5.2$ for $\epsilon = 10^{-7}$ as in (2), we find

* dbm = decibels relative to one milliwatt.

$$\begin{aligned}\bar{M}_\epsilon &= 18.6 + 23.7 - \bar{L} - 12 - 48 + m - \sigma - 10 \log (n/25) \\ &\quad - 20 \log 5.2 \\ &= m - \sigma - \bar{L} - 10 \log (n/25) - 32.\end{aligned}\tag{5}$$

Also,

$$\begin{aligned}\sigma_{M_\epsilon}^2 &= \sigma_S^2 + \sigma_L^2 + \sigma_A^2 + \sigma_{P_\epsilon}^2 \\ &= (0.5)^2 + (0.6)^2 + (1.0)^2 + (3.2)^2\end{aligned}$$

or

$$\sigma_{M_\epsilon} = 3.4 \text{ db},\tag{6}$$

the standard deviation of M_ϵ . The random variable M_ϵ is a measure of system performance which depends upon the statistical variations of signal level, pair loss, required amplifier signal-to-noise ratio, and near-end crosstalk interference. If, for a particular choice of the variables on the right side of (4), M_ϵ is positive, the amplifier operates at an error rate less than 10^{-7} ; if $M_\epsilon = 0$, the rate is 10^{-7} ; and, if M_ϵ is negative, the rate exceeds 10^{-7} . Assuming that M_ϵ has a normal distribution, 99 per cent of the amplifiers will operate with error rate less than $\epsilon (= 10^{-7}$ here) provided that

$$\bar{M}_\epsilon - 2.33\sigma_{M_\epsilon} = \bar{M}_\epsilon - 8.0 \geq 0\tag{7}$$

using (6). If we have equality in (7), i.e., if we place the one per cent point of the cumulative distribution of M_ϵ just at zero db, we will meet the error rate objective. Then from (5) and (7), rearranging, we have

$$\bar{L} + 10 \log (n/25) = m - \sigma - 40.\tag{8}$$

This equation gives upper bounds on both \bar{L} and n . That is, the mean section loss, \bar{L} , and the number of systems, n , must both be kept low enough so that the left side of (8) does not exceed the right side. If it does, we violate the error rate objective. Conversely, we may regard (8) as placing a lower bound on $m - \sigma$, derived from the properties of the cable, for a given \bar{L} and n .

As an example of the use of (8), assume $n = 200$ and 22-gauge unit cable, with pairs for the two transmission directions in adjacent splicing groups. Although Table II gives $m = 90$ db and $\sigma = 9$ db, we shall use the more conservative value $m = 84$ db. With this 6-db allowance we recognize that crosstalk measurements at 772 kc have been made on only a few of the many sizes and types of cable in the plant, and that some cables may have lower crosstalk losses than given in the

table. For $n = 200$ systems, (8) shows that \bar{L} must not exceed 26 db. If we accept this value of loss at 100 F, the loss at 55 F is 25 db. The latter figure, then, is the *design section loss* for this type of 22-gauge cable. That is, it is the maximum loss allowable for one-cable operation with 200 systems. For an engineering loss (Section 2.2) of 27 db per mile, a length of 4900 feet has 25-db loss, so this is the maximum distance allowable between repeater points for such an installation.*

We do not have data comparable to Table II for cable gauges other than 22. To estimate the near-end crosstalk losses for other gauges, we note that the near-end crosstalk coupling ratio varies, for the lengths and frequencies involved here, directly with the attenuation in db (Appendix A). For example, to estimate m for 19-gauge cable with average loss 19.1 db per mile, as in Table I, we use the corresponding m in Table II, and add $10 \log (19.1/26.5) = -1.4$ db. For 19-gauge layer cable, adjacent splicing groups, then, $m = 83 - 1.4 = 81.6$ db. Lacking more specific information, we may assume the standard deviation figures in Table II to hold for other gauges.

A value of \bar{L} from (8) which leads to a design section loss greater than 32.2 db simply means that the section loss may be 32.2 db, as we cannot exceed this value in any circumstances (Section 2.2). Using (8), with this limitation and with the correction for gauge described above, engineering tables have been prepared that give the design section losses for the various types of cable to be found in the field, for any number of systems.

4.2.2 Selection of Cables and Pairs

We would like to be free to assign systems to cable pairs as they are available along spans, without selecting individual pairs within the cable sections. Such assignment is possible with voice-frequency circuits, but T1 systems need more careful consideration. For one-cable operation, the two directions of transmission must be assigned to groups of pairs with a satisfactory range of near-end crosstalk coupling losses. More exactly, we wish to be sure that the mean loss and the standard deviation of losses are in accordance with Table II, or better,† in designing sections according to (8).

In large cables of unit construction, pairs are stranded into units of

* Amplifiers in one-cable sections are also exposed to far-end crosstalk from coupling paths among the pairs in the same transmission direction. To simplify the exposition in this section, the far-end crosstalk interference power, which adds to that due to near-end crosstalk, has been neglected. The far-end power may be calculated using the results of Section 4.3. When this is done, the result is equivalent to reducing the 6-db allowance slightly.

† I.e., a greater mean loss, or smaller standard deviation.

25 or 26 pairs for 19-gauge, 50 or 51 pairs for 22-gauge, and 100 or 101 pairs for 24- and 26-gauge cables. The units are identified by position and color coding. Measurements¹³ on a 900-pair, 22-gauge cable with paper pulp insulation, as in Fig. 9, show that the mean near-end loss at 750 kc for pairs in adjacent units, e.g., one pair in unit 1, the other in unit 2 of Fig. 9, is 13 db greater than for pairs within the same unit. The mean loss for pairs in alternate units, such as units 1 and 3 of Fig. 9, is 30 db greater than the within-unit value, and even greater differences are possible for pairs in, say, diametrically opposite units of a cable of this size. The standard deviations of the losses are not so dramatically changed by pair location, but are usually smaller for the more widely separated units. Similar remarks apply to cables of layer construction.

The common splicing group in large multipair cables is 100 pairs, i.e., two 50-pair units or four 25-pair units. For orderly administration of the outside plant, T1 pairs are assigned by splicing groups. There are three principal ways in which pairs for the two directions of transmission may be chosen, as given in Table III. The table also contains examples of the three configurations that apply to the unit cable of Fig. 9. If pairs for both directions are in the same splicing group, the design section losses from (8) are undesirably low. Placing the two transmission directions in adjacent groups increases the near-end crosstalk loss, and a reasonable number of systems can be operated in the same sheath with full repeater spacing. Going over to alternate-group separation,

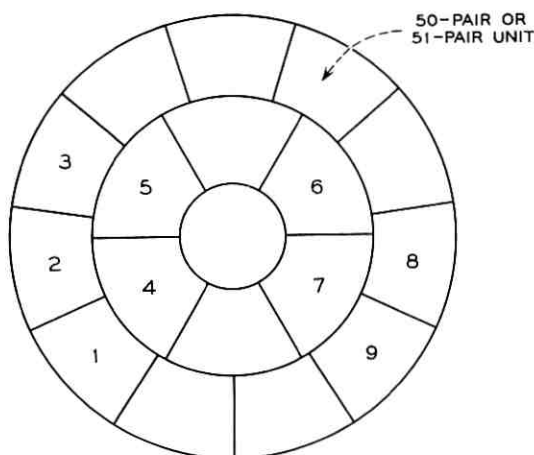


Fig. 9 — 900-pair, 22-gauge unit cable.

TABLE III—ASSIGNMENT OF SPLICING GROUPS IN ONE-CABLE OPERATION

Short Description	Explanation	Example for Unit Cable, Fig. 9		Design Section Loss in db, 22-gauge Unit Cable		
		E-W Pairs in Units	W-E Pairs in Units	25 Sys-tems	50 Sys-tems	100 Sys-tems
Same group	Pairs for both directions are in the same 100-pair splicing group.	1 & 2	1 & 2	24.0	21.1	18.2
Adjacent group	Pairs for the two directions are in separate groups that are next to one another.	1 & 2	4 & 5	32.2	30.7	27.9
Alternate group	Pairs for the two directions are in separate groups that have no units or layers touching.	1 & 2	6 & 7 or 8 & 9 etc.	32.2	32.2	32.2

there is a further increase in near-end crosstalk loss, and in the allowable number of systems. Therefore, to accommodate more systems in a given sheath, we may reduce repeater spacings or assign the systems to groups of pairs with greater separation.

It is not always possible, merely by examination of the usual outside plant records, to be sure that pair separation is maintained over a span length. In specific instances, especially where older cables are involved, it may be necessary to open some splices to determine the splicing pattern and to make rearrangements. If enough measurements for statistical validity can be made, it may be helpful to measure the near-end losses at or near 772 kc. In placing new cables where it is likely that T1 systems will be installed, modern splicing practices will be followed, including complete color-coded pair-to-pair splicing, as is possible with polyethylene-insulated cables.¹¹ This will generally make rearrangements or measurements unnecessary.

4.2.3 Staggered-Repeater Operation

Situations arise in planning T1 installations where one-cable operation is needed, but the crosstalk properties of the cable are not up to the requirements outlined above. For example, it may be doubtful if the pairs are in separate 100-pair splicing groups and inconvenient to open up enough splices to find out. In such cases, *staggered-repeater*

operation, as shown in Fig. 10, may be possible. Pairs for both transmission directions are in one cable, but the repeater cases are installed at staggered points so as to reduce near-end crosstalk interference. For example, the outputs of the amplifiers at repeater location A no longer face the inputs of those at B directly through the near-end crosstalk paths, but are first attenuated by L_1 db, the loss between the two locations. In effect, the mean near-end crosstalk coupling loss is increased by L_1 , and we are able to meet the interference objective, even though the same cable is unsuitable for one-cable operation as in Fig. 2(a).

The optimum layout for staggered-repeater operation is that in which all losses L_1, L_2, L_3 , etc. are exactly the same and are equal to half the design section loss. In a practical layout, it may not be possible to locate repeater cases in this way. Suppose the design section loss for one-cable use, without repeater staggering, is D db. If we require that each loss L_1, L_2 , etc. (Fig. 10) be at least 10 db, then we have a minimum near-end loss improvement of 10 db. The design loss D_s , between amplifiers in the same direction, is thus $D + 10$, or, if this number exceeds 32.2, $D_s = 32.2$. Of course, L_1, L_2 , etc. must each be less than $D_s - 10$.

At a constant value of design section loss, more systems may be installed in a given cable with staggered-repeater operation than in the usual one-cable mode.

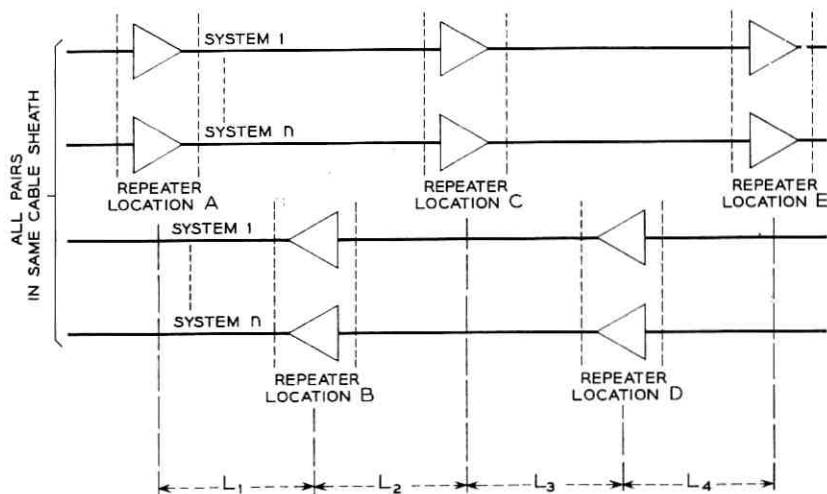


Fig. 10 — Staggered-repeater operation.

4.3 Two-Cable Systems

The general far-end crosstalk exposure of Fig. 11 is analyzed in Appendix B. The mean interfering power, \bar{Q}_n , at the preamplifier output of amplifier 2, due to crosstalk from amplifier 1, is given by (47), and the standard deviation of this power, $\sigma_{Q_n} = \sigma_{j,n}$ by (44). By selecting the values of L, L_1 and L_2 , we may determine limiting conditions for various layouts.

For example, let $L = L_1 = 32.2$ db, and $L_2 = 0$. This is the far-end coupling arrangement in an ordinary repeater section of maximum loss, with repeater locations aligned. Let us further assume that we have paper-insulated, 22-gauge cable with the properties listed in Table IV (Appendix B). As a worst case, suppose that there are 50 systems in the same cable unit, all but one of which may be regarded as interfering with the remaining one, or, closely, $n = 50$. The mean interfering power \bar{Q}_{50} , and its standard deviation, $\sigma_{Q_{50}}$, are calculated in Appendix B as -33.2 dbm and 5.2 dbm respectively.

Our calculation of error rate due to this interfering power is quite similar to that for near-end crosstalk, Section 4.2.1. As in (1), the power level exceeded by the interfering voltage with probability $\epsilon = 10^{-7}$, assuming again that the interference has a Gaussian amplitude distribution, is $P_\epsilon = Q_n + 20 \log 5.2 = Q_n + 14.3$. The mean over various selections of n crosstalk couplings is thus $\bar{P}_\epsilon = -33.2 + 14.3 = -18.9$ dbm. We may find the mean margin \bar{M}_ϵ by substituting in (4) the values $\bar{S} = 18.6, G = 23.7, L = 32.2, \bar{A} = 12$, and $\bar{P}_\epsilon = -18.9$; the result is $\bar{M}_\epsilon = 17.0$ db. For the standard deviation of M_ϵ , we take only that due to P_ϵ , or 5.2 db, as contributions from the variances of S and A are small. Thus 99 per cent of the amplifiers will have a margin at least

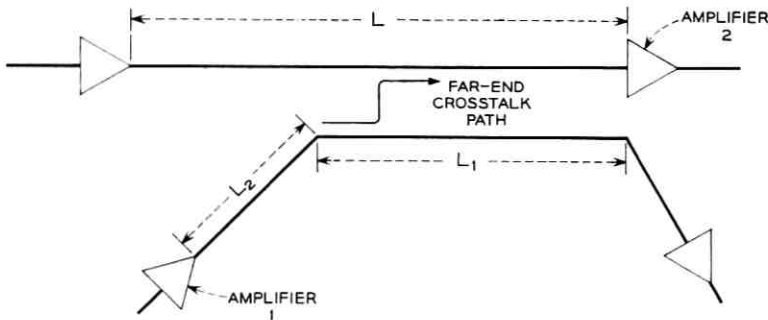


Fig. 11 — General far-end crosstalk exposure.

$\bar{M}_\epsilon - 2.33 \times 5.2 = 4.9$ db, again assuming a Gaussian distribution for M_ϵ .

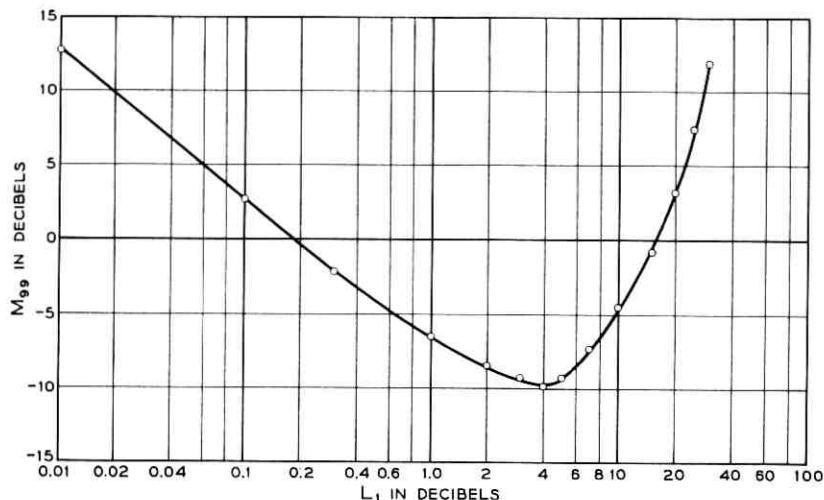
The effect of far-end crosstalk is more severe for an equivalent exposure in 19-gauge cable of lower loss per mile. Specifically, if we assume that the cable loss is 15.9 db per mile, and also that its crosstalk properties are the same as those used for the 22-gauge cable, the 99 per cent margin is 2.9 db for a 32.2-db section.

Neither of these values is quite as favorable as the 6-db reserve margin called for in our near-end crosstalk calculations (Section 4.2.1). As the 32.2-db section of 19-gauge cable, more than two miles in length, is rather an extreme case, we conclude that the line loss of ordinary repeater sections for one- or two-cable operation is not seriously limited by far-end crosstalk.

The calculation above relates to far-end interference within a single group of 50 systems in one cable unit. A different situation is shown in Figs. 8(a) and (b). Here we have far-end crosstalk between two groups of systems in two-cable operation, when the repeater locations for the two groups are different. This may come about, for example, when the number of repeater cases at each location is limited by manhole space, for underground systems. In Figs. 8(a) and (b), we have far-end paths between one set of amplifier outputs, and the inputs of another set, separated by a line section of loss L_M . To apply the theory, we allow L_1 of Fig. 11, which corresponds to L_M of Figs. 8(a) and (b), to vary from small values to 32.2 db, and again find the 99 per cent margin, $M_{99} = \bar{M}_\epsilon - 2.33 \sigma_{j, n}$, for $L_2 = 0$ and $L = 32.2$ db. For purposes of illustration, we select the cable properties as follows. A 600-pair, 22-gauge unit cable may have 400 systems in outside units surrounding 200 systems in interior units, in two-cable operation. We therefore use $n = 400$ interfering systems. The other cable parameters are taken from Table IV, choosing the adjacent-unit value, $m_{j_0} = 77$ db. This is pessimistic, since many of the far-end crosstalk coupling paths are between more widely separated pairs, in this instance. When we calculate M_{99} on this basis, we find a variation with L_1 as shown in Fig. 12.

For low values of L_1 , Fig. 11, the output of amplifier 1 is coupled to the input of amplifier 2 by a short far-end crosstalk path. This corresponds to the situation of Fig. 8(a), with L_M small. We see from Fig. 12 that L_1 must be less than 0.05 db in order to have M_{99} at least 6 db. We conclude that splices of apparatus cases to the main cable should be within 10 feet. Greater distances are tolerable for numbers of interfering systems less than 400.

The curve of Fig. 12 clearly shows that higher values of L_1 less than

Fig. 12 — M_{99} versus L_1 .

24 db will not provide the necessary margin. For sections spaced as in Fig. 8(b), then, we find that the interfering amplifier output must be at least 24 db from the disturbed amplifier input. This requirement cannot be met for sections limited to a maximum loss of 32.2 db. The reason is that interference in the opposite direction, i.e. from the 200-system group into the 400-system group in our example, also requires a minimum spacing, which we can estimate to be about 3 db less, or 21 db. This leads to a 45-db section loss. We conclude that misalignment of sections as in Fig. 8(b) is ruled out. More lenient spacing requirements result for cables of higher loss per mile, but the amount by which sections may be offset is still controlled by far-end crosstalk.

4.4 System Junctions

In an extensive area with T1 systems connecting several central offices, junctions such as those in Fig. 13 will occur. In the figure, systems connect office A with office C, and C with B, but for simplicity no through systems from A to B are included. The controlling far-end crosstalk exposures are in the cable sections incoming to amplifiers in office C, i.e., in cable section 3 for the two-cable case, Fig. 13(a), and in the incoming pairs of cable section 9 for the one-cable case, Fig. 13(b).

Assume in Fig. 13(c) that $L_B > L_A$, and that there are n interfering systems in the A branch which join systems from the B branch at the

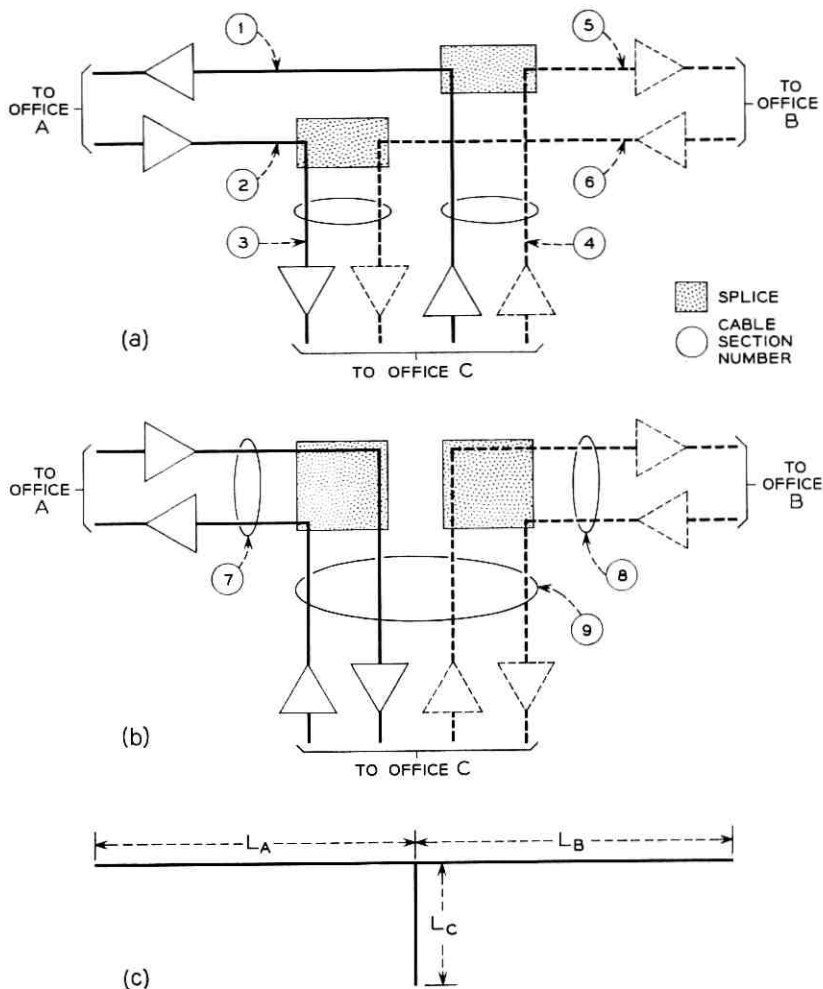


Fig. 13 — System junctions: (a) junction of two-cable systems; (b) junction of one-cable systems; (c) schematic showing losses.

junction. The far-end exposure has loss L_C , which corresponds to L_1 , Fig. 11. For L in Fig. 11 we have $L_B + L_C$, the loss in the signal path of the disturbed systems. For L_2 of Fig. 11 we have L_A , the loss in the disturbing pairs between the disturbing amplifiers and the junction.

The 99 per cent margin M_{99} is again given by

$$M_{99} = \bar{M}_\epsilon - 2.33 \sigma_{M_\epsilon} = \bar{M}_\epsilon - 2.33 \sigma_{j, n}$$

where j is the number of cable reels in the exposure, or C , section. We may find \bar{M}_ϵ as in (4), with $\bar{S} = 18.6$, $G = 23.7$, $L = L_B + L_C$, and $\bar{P}_\epsilon = \bar{Q}_n + 14.3$ as before. The necessary value of \bar{Q}_n , the mean far-end disturbing power in dbm, is obtained from (47) for $j = 1, 2, 3, \dots$ reels, or from (48) when the exposure section is shorter than one reel in length. When the required substitutions are made, again using the data of Table IV, and defining Δ as the maximum loss difference $L_B - L_A$ which will keep M_{99} at least 6 db, we find that Δ varies with L_C , the exposure section loss, as in Fig. 14. The number of systems, n , is a parameter in this figure. The curves are plotted on the assumption that systems entering the junction from the A side have at least adjacent-unit separation in the exposure section from systems entering from the B side, i.e. a mean equal-level far-end crosstalk loss at 772 kc, m_{j0} , of 77 db is used.

The value of Δ from Fig. 14 is the maximum permitted when far-end crosstalk is controlling, e.g., for an exposure as in cable section 3, Fig. 13(a). Other limits must also be taken into account. For example, suppose $n = 200$ systems are in the A branch, and the exposure length is $L_C = 10$ db. The value of Δ from Fig. 14 is 13.4 db. If $L_A = 12$ db it would seem that L_B might be as much as $12 + 13.4 = 25.4$ db. But the

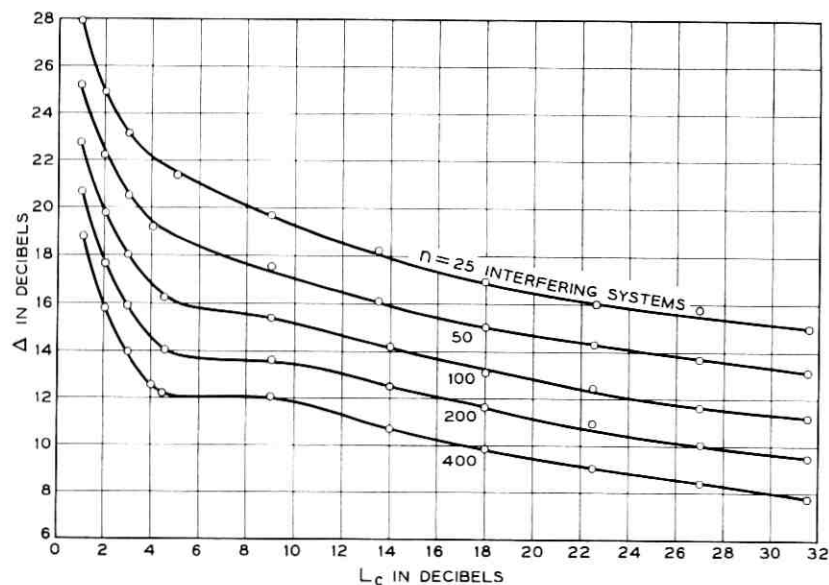


Fig. 14 — Δ versus L_C , 22-gauge cable.

total loss of the disturbed section, $L_B + L_C$, would then be $25.4 + 10 = 35.4$ db, more than the maximum possible loss of 32.2 db. For one-cable operation with design section loss less than 32.2 db, the restriction is based on this loss rather than 32.2 db; further, the margin M_{99} is affected by both near-end and far-end crosstalk interference power. Similarly, if the junction forms an entrance into a central office, the section losses $L_A + L_C$ and $L_B + L_C$ are limited as described in Section 4.5.

The curves of Fig. 14 may be used for mean far-end losses other than 77 db by adding $m_{j_0} - 77$ db to the indicated values of Δ , since Δ varies directly with this loss. For example, if there are 25 systems entering a junction of exposure length $L_C = 10$ db, and they are spliced into the same cable unit as the disturbed systems, we may estimate $m_{j_0} = 63$, so that $\Delta = 19.2 + 63 - 77 = 5.2$ db.

4.5 Sections Near Central Offices

Repeater sections that terminate in central offices, and others in which switched telephone pairs share a cable sheath with the T1 lines, are subject to office noise interference (Section 3.4). The fundamental information for engineering these sections is in Appendix C. As shown there, the signal-to-noise ratio, and hence the error rate, of an amplifier is determined by the difference between two losses. The first of these is the loss in the path by which the noise reaches the amplifier input. It is convenient to imagine that the noise originates at the office termination of the switched voice-frequency pairs. Thus, the noise path loss is a sum, in general, of a crosstalk coupling loss and a direct transmission loss. For the latter, it is adequate to substitute the loss at 772 kc of the pairs over which the noise travels. The second loss we need is that in the signal path, exclusive of the buildout network loss (Section 2.2), office pad loss (Section 2.3), and office cable loss. This is simply the cable loss incoming to the amplifier. In Fig. 20, error rate is plotted as a function of the difference in noise and signal path losses. An error rate of 10^{-7} is attained if this difference is at least 52 db. As explained in Appendix C, the error rates of Fig. 20 prevail under very severe noise conditions, so that we adopt this objective.

Fig. 15(a) shows the incoming T1 pair in a simple entrance section. For the office amplifier, the signal path loss is L_1 , the pair loss between the nearest outside amplifier and the main frame. The noise path loss, for the near-end crosstalk coupling path A, may be taken as 75 db (see Appendix C). We therefore must have $75 - L_1 \geq 52$, or $L_1 \leq 23$ db. For the first outside amplifier, the signal path loss is L_2 , and the noise

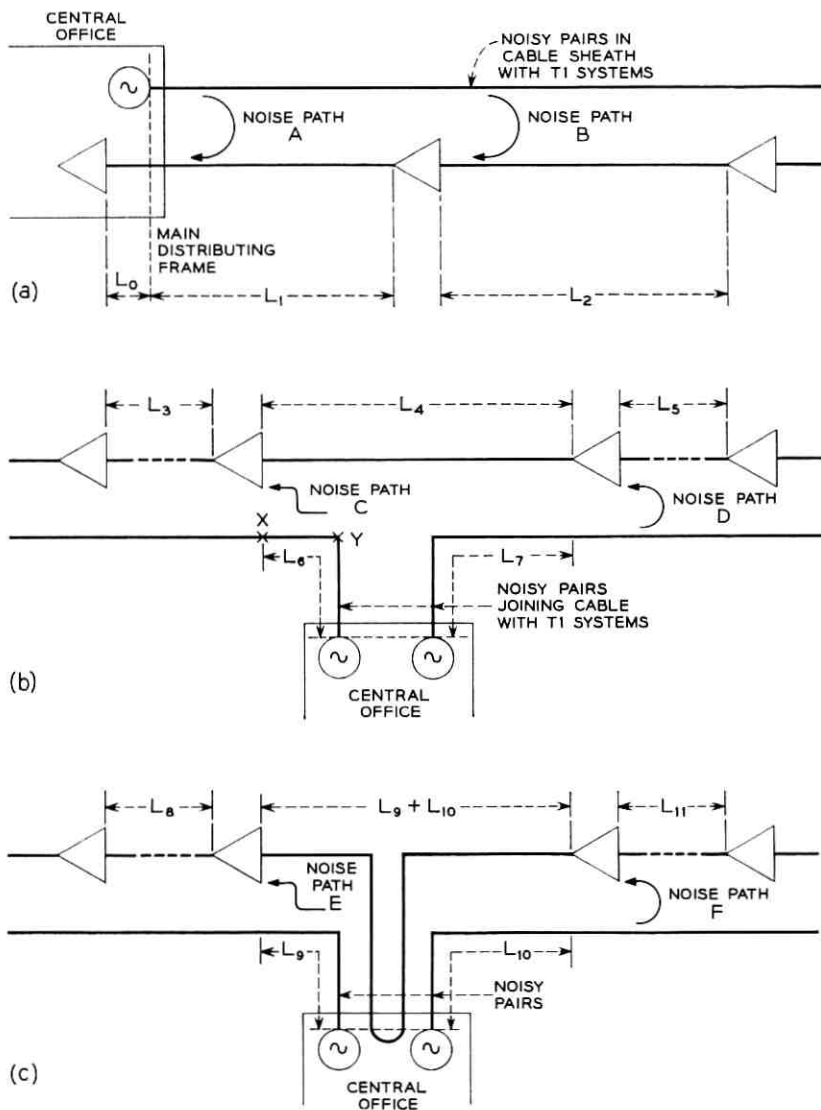


Fig. 15 — Repeater sections near central office: (a) T1 systems entering office; (b) T1 systems not entering office; (c) T1 pairs cross-connected at office main distributing frame.

path loss is $L_1 + 75$, for the near-end path B. We thus have $L_1 + 75 - L_2 \geq 52$, or $L_2 \leq L_1 + 23$. Notice that L_1 must be less than 9.2 db before any reduction in L_2 is needed, as L_2 cannot exceed 32.2 db. Sections farther from the office are not affected, owing to the increasing loss in the noise path.

In Fig. 15(b), one direction of T1 transmission is shown in an arrangement where the T1 lines do not enter the office, but are exposed to disturbing pairs which do enter it. The amplifier operating with signal loss L_4 has a total noise path loss, for path C, of $L_6 + \bar{L}_f$, where \bar{L}_f is the mean equal-level far-end crosstalk coupling loss at 772 kc for the cable section between points X and Y in the figure. [The distance from X to Y in feet may be substituted for l in (49), to approximate \bar{L}_f .] The losses L_4 and L_6 must satisfy the relation $L_6 + \bar{L}_f - L_4 \geq 52$. For near-end path D, we have a noise path loss of $L_7 + 75$, for the amplifier that operates with signal loss L_5 , so we require $L_7 + 75 - L_5 \geq 52$, or $L_5 \leq L_7 + 23$. For one-cable operation, with oppositely-directed amplifiers at the same points as those in Fig. 15(b), a similar set of inequalities applies, relating to noise paths not shown in the figure.

A third arrangement is that of Fig. 15(c), in which the T1 lines are cross-connected at the main frame, for flexibility in installing a terminal at the office later. For noise path E, we have $L_9 + \bar{L}_f - (L_9 + L_{10}) \geq 52$, or $\bar{L}_f - L_{10} \geq 52$; here \bar{L}_f is the far-end coupling loss for the entire L_9 section. For near-end path F, we find $L_{10} + 75 - L_{11} \geq 52$, or $L_{11} \leq 23 + L_{10}$. For one-cable operation, a comment similar to that above applies.

In arriving at the section limits due to office noise, we have not included a 6-db safety margin, as was done for near-end crosstalk in Section 4.2.1. The reason is apparent from Fig. 16, in which curves of error rate vs section loss are shown for (a) an end section, as in Fig. 15(a), designed with $L_1 = 23$ db, and (b) a one-cable section for which the design loss is 23 db. In the latter case, the error rate on the curve is that of the one per cent poorest amplifier; notice that this amplifier will not have an error rate greater than 10^{-7} until the section loss is 6 db more than the design section loss, or 29 db. The error rate due to crosstalk increases catastrophically beyond the 6 db margin. Since a section error rate of 10^{-6} is probably tolerable, while an error rate of 10^{-5} is not, there is as great a margin of safety for noise as for crosstalk. The steep slope of the curve of crosstalk error rate vs section loss in the critical error rate region necessitates considerable care in design.

In one-cable systems, the sections near offices are subject to interfer-

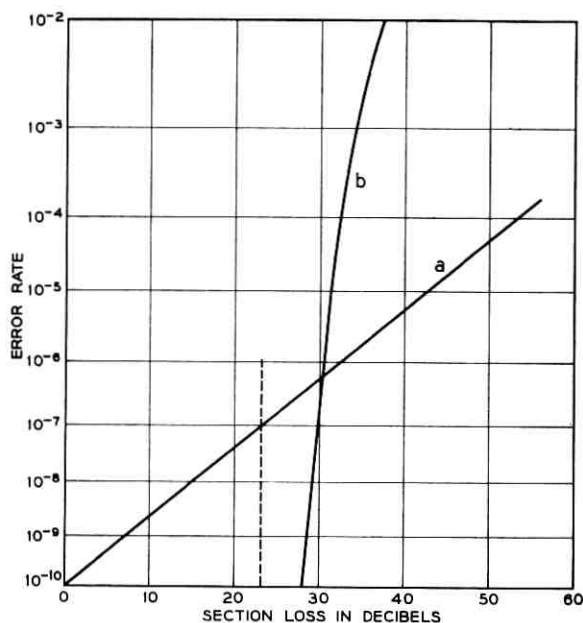


Fig. 16 — Error rates (a) in an office repeater due to switching noise and (b) in a one-cable section with design loss 23 db (poorest one per cent).

ence from near-end crosstalk, as discussed in Section 4.2, and also to office noise. For any such section, two losses may be found: the design section loss for one-cable operation (Section 4.2.1), and the maximum loss permitted by operation in the presence of central office noise. If we always choose the lesser of these two losses, we are assured of meeting the span error rate objective of Section 4.1. This simplifies the design, and we incur only a very small cost penalty in the number of repeaters needed.

The assignment of the T1 lines to cable pairs for either one- or two-cable operation may effectively isolate them from office noise disturbances, and the restrictions above may be relaxed. For example, in Fig. 15(a) the carrier pairs may be part of a splicing group in which the non-carrier pairs do not enter the office, but are cross-connected to outgoing pairs at the main frame. The prime source of office noise in the carrier systems is then coupling from noisy pairs in the same sheath which do enter the office, but are found in a different splicing group, with correspondingly greater crosstalk coupling losses to the carrier pairs. A

similar situation occurs when an entire splicing group, say a 50-pair unit, is initially equipped for T1, and the integrity of the group is maintained by unit-to-unit splicing, at least in the repeater sections near the office. Still another favorable case, for a layout like Fig. 15(b), is that in which the T1 pairs are in a group containing pairs spliced to an entrance cable, but to a complement of the entrance cable that does not enter the office. As described in Appendix C, noise measurements in these more favorable situations have shown that the effective crosstalk coupling losses in the noise paths are increased by 10 db. For example, in Fig. 15(a), we would use a loss of 85 db for near-end path A, instead of 75 db. As a result, there is no restriction arising from office noise, since the calculated maximum section loss exceeds 32.2 db.

Recalling the minimum line loss requirement for end sections of 6 db (Section 2.3), we find that these sections will lie in the range of 6 to 23 db, or, for 22-gauge cable, 1200 to 4500 feet. The advantage against office noise is great enough that the carrier wiring inside the office need not be physically separated from other wiring, e.g., by special cable racks or hangers. In most installations, the noise exposure inside will not be as great as that for which we have made allowance, by shortening the sections, in the outside plant.

V. CONCLUSION

In Sections II, III, and IV, we have given in piecemeal fashion the limitations on line layout that arise from the amplifier properties, and those that are needed to meet the error rate objective (Section 1.2.2) in the presence of interference. We conclude with a general description of how a system layout is made, taking into account these limitations.

Let us assume that the local trunk forecasts have shown the numbers of circuits needed and their allotment among various interoffice routes. These routes define the T1 spans. As far as possible, forecasts are made for a long growth period; each span must be designed for the ultimate number of T1 lines it will contain, rather than for the initial number of systems installed. A survey of the cable maps and circuit assignment records will show what cables and pairs are available for carrier use. With the present 25-repeater case, counts of 50 pairs must be cleared in each cable, for a minimum of 50 systems in two-cable operation or 25 systems in one-cable operation.* At this point, the decision as to the type of operation for each span may be made. When cables with assured

* As mentioned in Section 1.2.1, pairs may be equipped with through connectors or load coils instead of repeaters, when they are not needed immediately for carrier systems.

splicing characteristics (Section 4.2.2) are available, one-cable operation is attractive. Otherwise, and especially when very large numbers of systems are to be placed in a span, two-cable operation may be used. Splicing rearrangements and installation of new cable particularly suitable for carrier may also be considered at this stage.

A preliminary layout of each span is now made. Tentative repeater locations are selected so that repeater section lengths fall within the basic limits of 32.2 db, for two-cable operation, or the design section loss (Section 4.2.1) for one-cable. For simple spans directly between offices, section losses near the offices may readily be kept within the minimum and maximum limits of Section 4.5. Some estimates of office cabling losses are made at this point, as the office equipment bays are not yet installed. In two-cable operation, the outgoing section near an office is not limited in the same way as the incoming section, and they may be designed independently. It is quite possible to have the two directions in two-cable systems in different types of cable, or even along different routes.

When junctions occur, some revision of the preliminary layout may be called for, to meet the limits of Section 4.4. A common situation is a junction of two systems near an office, so that the section carrying inputs from both systems is subject to the end-section limits (Section 4.5) as well. Usually there is enough latitude to find a satisfactory layout without changing the fundamental plan. In exceptional cases, staggered-repeater operation (Section 4.2.3) may provide an alternative.

During the early stages of planning, a detailed examination of the cable routes, including manholes (for underground cable) and pole lines (for aerial cable) is essential. The results will show if some repeater locations must be shifted for traffic or other practical reasons. Further, it may be necessary to enlarge some manholes, or build auxiliary manholes, when many repeater cases are to be installed. Again, the necessary layout changes can most often be made well within the design limits.

For the convenience of operating telephone company engineers, the theoretical material developed here has been summarized in tables, e.g. for design section loss (Section 4.2.1). Along with the tables, general procedures for line layout, similar to those above, are given. Experience so far has shown that this information enables these engineers to plan line layouts in a simple and straightforward way, and that the resulting lines meet the transmission objective of Section 1.2.2. Unusual situations will arise in the field, as with any new system, that are not covered by the engineering information already supplied. As more experience is had with such situations, the theory here can be applied to them, and additional experimental data collected, with the goal of making the tables and recommendations more complete.

VI. ACKNOWLEDGMENTS

The experimental arrangements for the study of the effects of multi-system crosstalk on T1 repeaters were set up by W. L. Ross with the assistance of H. S. Piper, Jr., B. W. Boxx, and of D. B. Robinson, Jr., who designed the pulse generators. The experimental study itself was conducted largely by Robinson and R. G. DeWitt. An investigation of office noise which continued over several years involved many persons, but it is appropriate to mention the work of Robinson and of R. A. Gustafson, Jr. In numerous difficulties, we have had the assistance of K. E. Fultz, under whose direction all of this work was done, and to whom several basic concepts are attributable.

Much essential information about the existing telephone cable plant could only have been obtained through discussions with C. S. Thaeler and J. Mallett and others of the Laboratories and with engineers of the American Telephone and Telegraph Company. Finally, it is a pleasure to recall the discussions we have had of analytical methods with S. O. Rice.

APPENDIX A

Near-End Crosstalk

Consider two pairs in a cable section of length l and propagation constant γ , terminated in the characteristic impedance Z_0 at each end. A current I_0 is applied to one pair at one end. Let the mutual impedance unbalance at a distance x from this end be $Z(x)$ per unit length and the admittance unbalance be $Y(x)$. Then it can be shown¹⁶ that the incremental crosstalk current dI on the disturbed pair at the transmitting end, due to an incremental length of cable dx at distance x is given by

$$\frac{dI}{I_0} = - \left[\frac{Z_0 Y}{16} + \frac{Z}{4Z_0} \right] e^{-2\gamma x} dx. \quad (9)$$

The mutual impedance unbalance Z is due to inductance unbalance. The admittance unbalance Y is due primarily to capacitance unbalance. To a first-order approximation these effects are independent of frequency. At the frequencies of interest here Z_0 may be assumed to be a constant resistance. Equation (9) may therefore be written

$$\frac{dI}{I_0} = i\omega C(x) e^{-2\gamma x} dx \quad (10)$$

in which $C(x)$ is a real function of x , independent of frequency, which

will be called here the unbalance function.¹⁷ The statistical properties to be ascribed to $C(x)$ are as follows. Consider a large number of similar cables of length l . Each has, among others, two pairs, numbered 1 and 2, which are always similarly located with respect to each other. For each cable there is a $C(x)$ for these two pairs. The following assumptions are made for $C(x)$:

(i) When the length l is large enough, any kind of average for $C(x)$ over all values of x for one cable is equivalent to the same average taken at one value of x over all cables.

(ii) $C(x)$ is normally distributed in amplitude, with mean zero. Most of the following discussion does not depend on assumption (ii). In view of the first assumption, the autocorrelation function $S(r)$ and the power spectrum $G(s)$ may be defined as follows:

$$S(r) = \text{ave} [C(x)C(x + r)] \tag{11}$$

$$G(s) = 4 \int_0^\infty S(r) \cos 2\pi sr \, dr. \tag{12}$$

The response $R(\omega)$ is defined as the ratio of the total near-end crosstalk current to the current I_0 , and is obtained by integrating (10):

$$R(\omega) = i\omega \int_0^l C(x) e^{-2\gamma x} \, dx$$

or

$$R(\omega) = i\omega \int_0^l C(x) e^{-2\alpha x} (\cos 2\beta x + i \sin 2\beta x) \, dx \tag{13}$$

in which α and β are the attenuation and phase constants respectively. Since β is nearly proportional to frequency at video frequencies, $R(\omega)$ will fluctuate rapidly in a random way with frequency. The in-phase and quadrature components of $R(\omega)$ tend strongly to be normally distributed with mean zero regardless of the distribution of C because they are sums of many random components.

The "average" behavior of $R(\omega)$ is most easily examined by computing the average value of $|R|^2$. Denote this function by $p(\omega)$. Then

$$p(\omega) = \text{ave} \left\{ \omega^2 \int_0^l \int_0^l C(x)C(y) \exp [-2(\alpha + i\beta)x - 2(\alpha - i\beta)y] \, dy \, dx \right\}.$$

The averaging may be done inside the integral and (11) used to give

$$p(\omega) = \omega^2 \int_0^l \int_0^l S(x-y) \exp [-2\alpha(x+y) - 2i\beta(x-y)] dy dx. \quad (14)$$

Let $x + y = u$, $x - y = v$. Equation (14) becomes

$$p(\omega) = \frac{\omega^2}{2} \int_0^l \int_{-u}^u S(v) e^{-2\alpha u - 2i\beta v} dv du \\ + \frac{\omega^2}{2} \int_l^{2l} \int_{-2l+u}^{2l-u} S(v) e^{-2\alpha u - 2i\beta v} dv du.$$

Making use of the fact that $S(v)$ is even and changing variables in the second integral leads to

$$p(\omega) = \omega^2 \int_0^l (e^{-2\alpha u} + e^{2\alpha u - 4\alpha l}) \int_0^u S(v) \cos 2\beta v dv du.$$

This can be integrated by parts to give

$$p(\omega) = \frac{\omega^2}{2\alpha} \int_0^l (e^{-2\alpha u} - e^{2\alpha u - 4\alpha l}) S(u) \cos 2\beta u du.$$

It will be seen shortly that at video frequencies $S(u)$ is extremely small for values of u large enough to make $e^{-2\alpha u}$ substantially different from unity. Using this fact and (12), we have¹⁷ for l greater than a few feet,

$$p(\omega) = \frac{\omega^2}{8\alpha} (1 - e^{-4\alpha l}) G(\beta/\pi). \quad (15)$$

Over the video range β is approximately proportional to frequency and may be replaced by $2\pi f/c$ where c is the propagation velocity in miles per second. The length factor $1 - e^{-4\alpha l}$ is unity for frequencies and spacings prevalent in T1 applications so (15) becomes

$$p(\omega) = \frac{\omega^2}{8\alpha} G(2f/v). \quad (16)$$

The function $p(\omega)$ has been found empirically by averaging many cross-talk measurements to increase 15 db per decade over the range 100 kc to 10 mc (above the latter frequency its behavior is unknown). Therefore, since α is approximately proportional to the square root of ω , $G(2f/c)$ is nearly constant over this range. This yields some information concerning $S(r)$. This function is required to take its maximum value at the origin.

If we assume the function

$$S(r) = S(0)e^{-k_2|r|} \quad (17)$$

then G is given by

$$\begin{aligned} G(x) &= 4S(0) \int_0^{\infty} e^{-k_2 r} \cos 2\pi x r \, dr \\ &= \frac{4S(0)k_2}{4\pi^2 x^2 + k_2^2}. \end{aligned}$$

Consequently $G(2f/c)$ is given by

$$G(2f/c) = \frac{4S(0)k_2 c^2}{16\pi^2 f^2 + k_2 c^2}. \quad (18)$$

The departure of $p(\omega)$ from the 15-db per decade line is of the order of the measurement inaccuracy, about 1 db. The value of k_2 can be estimated by assuming $G(2f/c)$ falls off 1 db at 10 mc from its value at the origin. Assuming $c = 1.2 \times 10^8$ miles per second, this leads to the result

$$1/k_2 = 5 \times 10^{-4} \text{ miles}. \quad (19)$$

Consequently the correlation function $S(r)$ has decreased to e^{-1} times its value at the origin when r is about 2.6 feet; in other words, the crosstalk unbalances at points further apart than this are essentially uncorrelated. The object of staggering pair twists is to make this correlation range small, i.e., to increase k_2 , since as (18) shows, the crosstalk power is inversely proportional to k_2 in the frequency range of interest. It appears that further improvement in near-end crosstalk by this means will be difficult to achieve.

Having found the average behavior of the crosstalk function $|R(\omega)|$ with frequency, it is worth while to look at its amplitude distribution. For the ensemble of cables pictured earlier, the average value of $|R(\omega)|^2$ for a particular pair combination is given by (16), but a given cable will generally have a different value, and the distribution of these values enters into crosstalk calculations. The distribution of $|R(\omega)|$ may be investigated by dealing with the real and imaginary parts of $R(\omega)$, denoted by $X(\omega)$ and $Y(\omega)$, respectively. Let

$$\overline{X^2} = m_{11}, \quad \overline{Y^2} = m_{22}, \quad \overline{XY} = m_{12}.$$

Since X and Y are normally distributed, the joint probability distribution $p(x,y)$ of X and Y is given by

$$\begin{aligned} p(x,y) &= (2\pi)^{-1} (m_{11}m_{22} - m_{12}^2)^{-1} \\ &\quad \cdot \exp \left\{ \frac{-m_{22}x^2 - m_{11}y^2 + 2m_{12}xy}{2(m_{11}m_{22} - m_{12}^2)} \right\}. \quad (20) \end{aligned}$$

From this the desired distribution of $|R| = (X^2 + Y^2)^{1/2}$ can be obtained. From (13), for a long cable section,

$$X = -\omega \int_0^{\infty} e^{-2\alpha x} C(x) \sin 2\beta x \, dx$$

$$Y = \omega \int_0^{\infty} e^{-2\alpha x} C(x) \cos 2\beta x \, dx.$$

Multiplication and averaging leads to the results

$$m_{11} = \frac{\omega^2}{2} \int_0^{\infty} e^{-2\alpha u} \cos 2\beta u \int_0^u S(v) \, dv \, du \\ + (4\alpha)^{-1} \omega^2 \int_0^{\infty} e^{-2\alpha u} S(u) \cos 2\beta u \, du$$

$$m_{22} = -\frac{\omega^2}{2} \int_0^{\infty} e^{-2\alpha u} \cos 2\beta u \int_0^u S(v) \, dv \, du \\ + (4\alpha)^{-1} \omega^2 \int_0^{\infty} e^{-2\alpha u} S(u) \cos 2\beta u \, du$$

$$m_{12} = \frac{\omega^2}{2} \int_0^{\infty} e^{-\alpha u} \sin 2\beta u \int_0^u S(v) \, dv \, du.$$

Substitution of an estimate of $S(r)$ such as is given by (17) and (19) shows that m_{11} is approximately equal to m_{22} and m_{12} may be neglected over the range 100 kc to 10 mc. Equation (20) then becomes

$$p(x,y) = (2\pi m_{11})^{-1} \exp \left[-\frac{x^2 + y^2}{2m_{11}} \right].$$

The probability that the point (X,Y) lies in the differential area $dx \, dy$ is given by $p(x,y) \, dx \, dy$. By making the change of variable $x = r \cos \theta$, $y = r \sin \theta$ the probability density of $|R|$ may be found to be

$$p_1(r) = \frac{r}{m_{11}} \exp [-r^2/2m_{11}], \quad (r \geq 0).$$

The phase angle θ is uniformly distributed over $(0, 2\pi)$.

The discussion of near-end crosstalk may be summarized as follows. For a given pair combination in a reel of cable which is long enough so that $e^{-4\alpha l}$ is small in comparison to unity, the magnitude of the near-end crosstalk response $R(\omega)$ to a one-volt sinusoid of frequency $\omega/2\pi$ has the probability density

$$\text{prob} (r < |R| < r + dr) = p_1(r) \, dr = \frac{2r}{p(\omega)} e^{-r^2/2m_{11}} \, dr. \quad (21)$$

Over the frequency range 100 kc to 10 mc, $p(\omega)$ for multipair cables is given quite accurately by

$$p(\omega) = k\omega^3 \quad (22)$$

in which k is a constant which varies from one combination of two pairs to another.

The statistical properties of the values of k corresponding to the many possible combinations of two pairs in a cable may not be predicted from this theory but must be estimated from measurements of crosstalk loss. The analysis up to this point permits only the prediction of the statistical properties of measurements made at different frequencies on a single combination of two pairs. Let $L(\omega)$ be the crosstalk loss in db for a particular pair combination at a frequency $\omega/2\pi$. Then $L(\omega)$ is a random function of frequency, with an amplitude distribution and therefore a mean and standard deviation at each frequency. It will be convenient in what follows to use the probability density of $|R|^2$ rather than of $|R|$; this is found from (21) to be

$$\text{prob} (u < |R|^2 < u + du) = p_2(u) = \frac{1}{p(\omega)} e^{-u/p(\omega)}.$$

For a given k corresponding to a particular choice of disturbing and disturbed pairs, the expected value of the crosstalk loss $L(\omega)$ in db as a function of frequency is given by

$$\begin{aligned} \langle L(\omega) \rangle &= -\langle 10 \log |R|^2 \rangle \\ &= -10 \log e \langle \ln |R|^2 \rangle. \end{aligned} \quad (23)$$

From (21),

$$\begin{aligned} \langle \ln |R|^2 \rangle &= \frac{1}{p(\omega)} \int_0^\infty \ln u e^{-u/p(\omega)} du \\ &= \psi(1) + \ln p(\omega) \end{aligned}$$

in which $\psi(x)$ is the derivative of the logarithm of the gamma function,^{18,19} and $\psi(1) = -0.577$. Using (22), (23) becomes

$$\langle L(\omega) \rangle = -10\psi(1) \log e - 10 \log k - 15 \log \omega. \quad (24)$$

The variance of $L(\omega)$ is given by

$$\langle L^2(\omega) \rangle - \langle L(\omega) \rangle^2 = 10^2 \psi'(1) \log^2 e$$

from which the standard deviation σ is found to be

$$\sigma = 10 \log e \sqrt{\psi'(1)} = 5.56 \text{ db}. \quad (25)$$

Equation (25) shows that the standard deviation of the loss in db is independent of frequency. If the same pair combination in many different cables is tested at a given frequency, or in other words if the loss in db is measured at one frequency on many pair combinations having the same value of k , the standard deviation of all the readings will be 5.56 db at any frequency in the range of 100 kc to 10 mc. The average value at any frequency ω will be given by (24).

In practice, crosstalk loss measurements have been made at relatively few frequencies, often at only one frequency which is of particular interest in a given application, but on many pair combinations. The mean value of the losses in db of all the pair combinations decreases 15 db per decade in the band from 100 kc to 10 mc. The standard deviation is always greater than 5.56 db, due to the contribution of the variation in the value of k from one pair combination to another. Measurements¹³ of near-end crosstalk loss in a 22-gauge cable of unit construction yielded standard deviations of about 9 db for pair combinations in which both pairs were in the same unit, 7.5 db for combinations in which the two pairs were in adjacent units, and 6.5 db for combinations in which the two pairs were separated by one unit. The distribution of the loss in db at any frequency fits the normal law quite well and the normal distribution is usually assumed in making calculations.

In order to complete a picture of near-end crosstalk which will be adequate for repeated line engineering, it is only necessary to find the distribution of the k 's which together with the distribution

$$p_2(u) = \frac{1}{k\omega^3} \exp - (u/k\omega^3)$$

will give the empirically obtained normal distribution of crosstalk losses. The latter distribution is assumed to have a mean loss of m db at a reference frequency ω_0 and a standard deviation of σ db at all frequencies. At an arbitrary frequency ω , the probability density of loss $L(\omega)$ in db is given by

$$\begin{aligned} \text{prob } (x < L(\omega) < x + dx) \\ = \frac{1}{\sqrt{2\pi}\sigma} \exp - \left[\left(x - m - 15 \log \frac{\omega}{\omega_0} \right)^2 / 2\sigma^2 \right] dx. \end{aligned}$$

Since $L(\omega) = -10 \log |R(\omega)|^2$, the density of $|R|^2$ is given by

$$\begin{aligned} \text{prob } (x < |R|^2 < x + dx) \\ = \frac{10 \log e}{\sqrt{2\pi}\sigma x} \exp - \left[\left(10 \log x + m - 15 \log \frac{\omega}{\omega_0} \right)^2 / 2\sigma^2 \right] dx. \end{aligned}$$

If we set $C = (10 \log e)^{-1}$ the density becomes

$$\frac{1}{\sqrt{2\pi C'\sigma x}} \exp - \left[\left(\ln x + mC - \frac{3}{2} \ln \frac{\omega}{\omega_0} \right)^2 / 2C^2\sigma^2 \right] dx.$$

It is convenient to redefine k in (22) so that

$$p(\omega) = k \left(\frac{\omega}{\omega_0} \right)^{\frac{3}{2}}. \tag{26}$$

Now what is needed is a probability density $p_k(x)$ for k such that

$$\begin{aligned} \frac{1}{\sqrt{2\pi C'\sigma x}} \exp - \left[\left(\ln x + mC - \frac{3}{2} \ln \frac{\omega}{\omega_0} \right)^2 / 2C^2\sigma^2 \right] \\ = \int_0^\infty p_k(u) \frac{1}{u(\omega/\omega_0)^{\frac{3}{2}}} \exp - \left[\left(\frac{x}{u} \right) / \left(\frac{\omega}{\omega_0} \right)^{\frac{3}{2}} \right] du. \end{aligned}$$

However, an accurate solution of this equation is not necessary. The lognormal distribution itself is only a convenient approximation to the actual distribution. For p_k we require a function which is a probability density, whose convolution may be expressed in elementary functions, which possesses at least two arbitrary parameters, and which, applied to the right side of (24), permits the calculation of the first and second moments of $\langle L(\omega) \rangle$. A satisfactory function is

$$p_k(x) = \frac{\lambda^\nu}{\Gamma(\nu)} x^{\nu-1} e^{-\lambda x} (\nu > 0, \lambda \geq 0).$$

If we now redefine k in (24) to agree with (26), and denote expectations relative to $p_k(x)$ by E , then

$$\begin{aligned} E\langle L(\omega) \rangle &= m - 15 \log \left(\frac{\omega}{\omega_0} \right) \\ &= -10 \log e [\psi(1) + \psi(\nu)] + 10 \log \lambda - 15 \log \left(\frac{\omega}{\omega_0} \right). \end{aligned}$$

Similarly the variance in db may be calculated for each distribution to yield

$$E(\langle L(\omega) \rangle)^2 - (E\langle L(\omega) \rangle)^2 = (10 \log e)^2 [\psi'(1) + \psi'(\nu)] \sigma^2.$$

Consequently we set

$$\begin{aligned} \psi'(\nu) &= (10 \log e)^{-2} \sigma^2 - \psi'(1) \\ 10 \log \lambda &= m + 10 \log e [\psi(1) + \psi(\nu)]. \end{aligned} \tag{27}$$

The results so far may be summarized as follows. Suppose the dis-

tribution of near-end crosstalk losses between two groups of pairs in a given type of exchange cable is found empirically to possess a mean of m db and a standard deviation of σ db at a frequency ω_0 . Then for each combination consisting of a pair in each group, the probability density of $|R(\omega)|^2$, the squared amplitude ratio of crosstalk voltage to transmitted voltage, is

$$\text{prob}(x < |R(\omega)|^2 < x + dx) = k^{-1} \left(\frac{\omega}{\omega_0} \right)^{-3} \exp \left[-xk^{-1} \left(\frac{\omega}{\omega_0} \right)^{-3} \right] dx. \quad (28)$$

The distribution of the k 's for all the pair combinations is given by

$$\text{prob}(x < k < x + dx) = p_k(x) dx = \frac{\lambda^\nu}{\Gamma(\nu)} x^{\nu-1} e^{-\lambda x} dx \quad (29)$$

in which λ and ν are given by (27). In (28), the average value of $|R(\omega)|^2$ is $k(\omega/\omega_0)^3$.

We may now calculate the mean and standard deviation of the interference power at the preamplifier output when there are n systems in operation. Let $r(f)$ be the Fourier transform of the preamplifier impulse response, and $W(f)$ the average power spectrum of the pulse train at the amplifier output in watts per cycle. Consider a hypothetical interferer coupled to the disturbed amplifier by a near-end crosstalk path whose power ratio at frequency f is just $k(f/f_0)^3$. If the expected value of the interference power due to this interferer is denoted by $w_c(k)$, then

$$w_c(k) = k \int_0^\infty |r(f)|^2 W(f) \left(\frac{f}{f_0} \right)^3 df.$$

The actual interference power will be close to this value because in summing $|R(\omega)|^2$ over a range of frequencies the variation tends to be averaged out. The standard deviation of w_c for a given value of k is taken to be zero.

The total interference power is the sum of many contributions $w_c(k)$ with k 's chosen at random from the distribution given by (29), in which λ and ν are determined from the mean m and standard deviation σ in db of the crosstalk loss distribution. The distribution of interference power due to n interferers depends upon the distribution of the random variable which is the sum of n values of k chosen at random, say k_n . This is the n th convolution $p_k^{(n)}(x)$ of $p_k(x)$ and is given by

$$p_k^{(n)}(x) = \frac{\lambda^{n\nu}}{\Gamma(n\nu)} x^{n\nu-1} e^{-\lambda x}. \quad (30)$$

The mean and standard deviation of the interference power w_c in dbm

at the preamplifier output may now be determined. From (30) one finds

$$\text{ave} (10 \log k_n) = 10 \log e [\psi(n\nu) - \ln \lambda]$$

$$\text{ave} (100 \log^2 k_n) - [\text{ave} (10 \log k_n)]^2 = (10 \log e)^2 \psi'(n\nu).$$

The mean interference power in dbm at the preamplifier output is therefore given by

$$\begin{aligned} \bar{P}_n &= \text{ave} [10 \log w_c(k_n)] + 30 \\ &= 30 + 10 \log e [\psi(n\nu) - \psi(\nu) - \psi(1)] \\ &\quad - m + 10 \log \left\{ \int_0^\infty |r(f)|^2 W(f) \left(\frac{f}{f_0}\right)^{\frac{3}{2}} df \right\}. \end{aligned}$$

The standard deviation is

$$\sigma_{P_n} = 10 \log e [\psi'(n\nu)]^{1/2}. \quad (31)$$

The value of ν must be determined from (27).

The quantity

$$30 + 10 \log \left\{ \int_0^\infty |r(f)|^2 W(f) \left(\frac{f}{f_0}\right)^{\frac{3}{2}} df \right\} - 10 \log e \psi(1) \quad (32)$$

is the interference in dbm at the preamplifier output due to one amplifier when $m = 0$ at frequency $f_0 = \omega_0/2\pi$. This frequency will be taken to be half the repetition frequency of the pulse train, or 772 kc. The preamplifier response in db is given in Fig. 17. The expression for $W(f)$ may

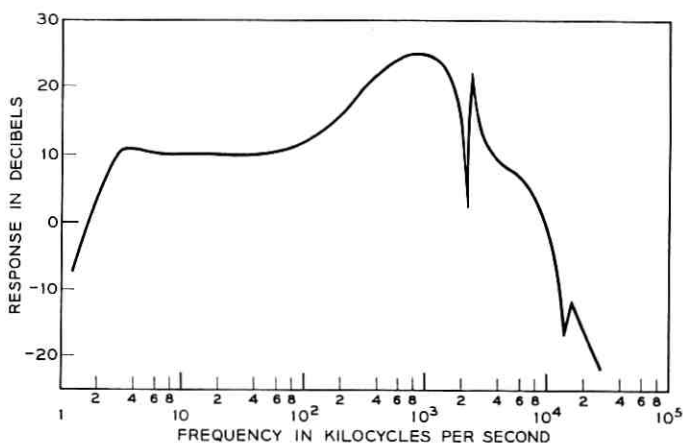


Fig. 17 — Preamplifier response.

be written⁴ as

$$W(f) = \frac{2V^2 f_0}{100\pi^2 f^2} \sin^2(\pi f/4f_0) \left(1 - \cos \frac{\pi f}{f_0}\right). \quad (33)$$

This assumes rectangular pulses of height V volts base-to-peak, $1/4f_0$ seconds long at an impedance level of 100 ohms, with probability $\frac{1}{2}$ for pulse and $\frac{1}{2}$ for space, each time slot being independent of preceding ones. It also takes into account the polarity constraint that each pulse is of opposite polarity from the preceding one; see Fig. 3. When $V = 3$ volts is substituted in (33), and (32) is integrated numerically, the result is 39.4 dbm. The mean interference power in dbm at the pre-amplifier output due to n interfering amplifiers is therefore

$$\bar{P}_n = 39.4 + 10 \log e [\psi(n\nu) - \psi(\nu)] - m$$

in which m is the mean crosstalk loss in db at 772 kc and ν is to be determined from the standard deviation σ of crosstalk loss in db by

$$\begin{aligned} \psi'(\nu) &= \sigma^2(10 \log e)^{-2} - \psi'(1) \\ &= 0.0530\sigma^2 - 1.645. \end{aligned}$$

For engineering calculations, a convenient approximation to \bar{P}_n is

$$\bar{P}_n = 48 - m + \sigma + 10 \log(n/25). \quad (34)$$

This equation, based on well-known properties of the ψ -function,^{18,19} is valid within 0.2 db for σ between 6 and 14 db.

These results for \bar{P}_n are in good agreement with experiment, but (31) has been found to yield too small a value for σ_{P_n} . This may be traced to the neglect of the variance of $w_c(k)$ and the use of its average value only. The exact analysis has not been attempted, but an experimental value¹⁵ of σ_{P_n} is used in engineering (Section 4.2.1).

A crucial point in the application in Section 4.2 of these results on near-end crosstalk is the Gaussian amplitude distribution of the interfering voltage. This was established experimentally¹⁵ with the setup of Fig. 6, as follows. A fixed crosstalk output was obtained by choice of the connections at the cross-connect terminal. The variable gain in the crosstalk path (see Fig. 6) was adjusted to cause an error rate of 1.2×10^{-6} , a convenient value for measurement, in the receiving T1 amplifier. The interfering power for this condition was taken as a reference value. Next, the interfering power was varied from the reference level, by changing the gain setting, and the variation of error rate observed. The resulting points, for 3 and 25 interferers, are plotted in Fig. 18. Quite similar results were found for 1, 6, 12, and 50 interferers.

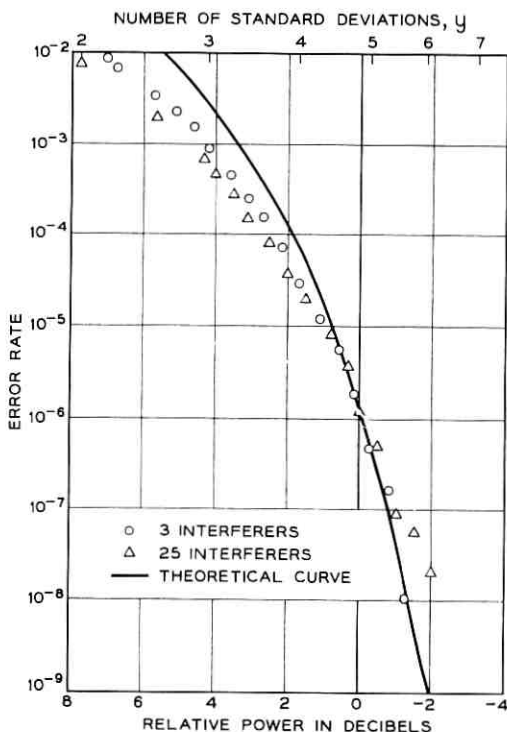


Fig. 18 — Error rate versus relative power.

The solid curve of Fig. 18 is the theoretical result that would be obtained if the crosstalk had a Gaussian amplitude distribution and a power level such as to cause an error rate of 1.2×10^{-6} in the reference condition. As a Gaussian signal exceeds an amplitude of 4.85 times its standard deviation (rms value) with probability* 1.2×10^{-6} , we have plotted $y = 4.85$ on the upper horizontal scale of Fig. 18, to coincide with 0 db on the lower scale. For a value of power G db greater than the reference, the value of y is $4.85 \times 10^{-G/20}$. For example, if the gain is set to give 2 db less power than in the reference condition, or $G = -2$, the Gaussian interference must exceed $4.85 \times 10^{2/20} = 6.1$ times its rms value to cause an error, an event with probability 1.1×10^{-9} .

Measurements of the near-end crosstalk amplitude distribution out to

* This is the "double-tail" probability, or twice the value ϵ in (2), for a given number of standard deviations, y , because the data were taken with a one-out-of-eight pulse pattern as the transmitted signal, and most errors are insertions of pulses. The probability of such errors is 2ϵ .

two or three standard deviations have been made with a conventional amplitude distribution analyzer. These measurements confirm the Gaussian distribution to a probability level of 10^{-3} or so. The confirmation we have described for the lower probabilities that occur for five or six standard deviations is essential in line engineering calculations with error rates as low as 10^{-7} .

APPENDIX B

Far-End Crosstalk

Consider again two pairs of propagation constant $\gamma(f)$ in a cable section of length l , which are terminated in their characteristic impedance Z_0 at each end. A current I_0 of frequency f is applied to one pair at one end. Let the mutual impedance unbalance at a distance x from this end be $Z(x)$ per unit length and the admittance unbalance be $Y(x)$. Then it can be shown¹⁶ that the incremental crosstalk current dI on the disturbed pair at the receiving end, due to an incremental length of cable dx at distance x , is given by

$$\frac{dI}{I_0} = \left(\frac{Z_0 Y}{16} - \frac{Z}{4Z_0} \right) e^{-\gamma l} dx. \quad (35)$$

At frequencies of interest we may make the approximation that Z , Z_0 , and Y are independent of frequency. Equation (35) may then be written

$$dI/I_0 = i\omega C_r(x) e^{-\gamma l} dx$$

in which $C_r(x)$ may be interpreted as the unbalance per unit length. It is convenient to deal with "equal-level" crosstalk loss, that is, with the ratio $dI/I_r = i\omega C_r(x) dx$ where I_r is the current on the disturbing pair at its receiving end. If we define $F(f)$ as the response of the disturbed pair at the receiving end of the cable to a sinusoid impressed on the disturbing pair at the sending end which is large enough to produce unit amplitude at the receiving end, then

$$F(f) = i2\pi f \int_0^l C_r(x) dx. \quad (36)$$

The quantity of immediate interest is the expected value of $|F|^2$. From (36), and setting $\langle C_r(x)C_r(y) \rangle = S_r(x - y)$ we get

$$\langle |F(f)|^2 \rangle = 2\omega^2 \int_0^l (l - x) S_r(x) dx.$$

If l is very much larger than the range over which $S_r(x)$ is non-negligible, the above equation becomes

$$\langle |F(f)|^2 \rangle = 2\omega^2 l \int_0^\infty S_r(x) dx. \tag{37}$$

This result shows the well-known 6-db per octave average slope of the response with frequency and the proportionality to cable length.¹⁷

From (36) we expect the function $i^{-1}F(f)$ to be normally distributed with mean zero since $C_r(x)$ has mean zero. This would be true even if C_r were not normally distributed, provided that l is large compared to the range over which C_r is self-correlated. Thus for a particular pair combination, $i^{-1}F(f)$ will have a value in the range $(x, x + dx)$ with probability

$$(2\pi)^{-\frac{1}{2}} \mu^{-1}(f) \exp \{ - [x^2/2\mu^2(f)] \} dx,$$

where μ is the standard deviation. From this the probability density of $|F(f)|^2$ is found to be

$$(2\pi x)^{-\frac{1}{2}} \mu^{-1} \exp - (x/2\mu^2).$$

The first moment of this distribution is $\mu^2(f)$, which may therefore be equated to the right-hand side of (37). It is convenient to set

$$2 \int_0^\infty S_r(x) dx = \frac{k}{\omega_0^2}.$$

Then $\mu^2(f) = kl\omega^2/\omega_0^2$ and kl is the expected value of the power on the disturbed pair at the receiving end, per watt of received power in the disturbing pair, at a frequency ω_0 . As in the near-end case, k is a random variable, a particular value being associated with each pair combination. We assume that the value of k possessed by a pair A in combination with pair B is independent of its value in combination with any other pair. For the probability that k for a pair combination chosen at random lies in the range $(u, u + du)$ we assume

$$P_k(u) = \frac{\lambda^\nu}{\Gamma(\nu)} u^{\nu-1} e^{-\lambda u} \quad (\nu > 0, \lambda \geq 0). \tag{38}$$

The probability that the equal-level far-end crosstalk response $|F(f)|^2$ of a pair combination chosen at random at a frequency f lies in the range $(x, x + dx)$ is therefore given by

$$P(x) dx = dx \frac{\lambda^\nu}{\Gamma(\nu)} \int_0^\infty u^{\nu-1} \frac{\omega_0}{\omega \sqrt{2\pi l u x}} \exp (-\lambda u - \omega_0^2 x / 2\omega^2 ul) du.$$

The equal-level far-end crosstalk loss in db is given by $-10 \log |F|^2$. For the above distribution we have the mean loss

$$m_1 = -(10 \log e) \left[\psi(\nu) + \psi\left(\frac{1}{2}\right) \right] + 10 \log \lambda + 10 \log \frac{\omega_0^2}{2\omega^2} - 10 \log l,$$

and the standard deviation in db

$$\sigma_1 = (10 \log e) \left[\psi'(\nu) + \psi'\left(\frac{1}{2}\right) \right]^{\frac{1}{2}}.$$

A repeater section is usually made up of several reels spliced together. For a given pair in a cable section of j reels the far-end coupling is therefore made up of j couplings in parallel, each having a value of k chosen at random from the distribution (38). Each coupling may be thought of as a lumped capacitance which may be positive or negative and is equal to the integral of $C_r(x)$. The currents therefore add linearly. The density of the resultant current, relative to the received current is, therefore, the j -fold convolution of the density function of $i^{-1}F(f)$, i.e.

$$\lambda^{\nu} [\Gamma(\nu)]^{-1} \int_0^{\infty} u^{\nu-1} (\omega_0/\omega\sqrt{2\pi lu}) \exp(-\lambda u - \omega_0^2 x^2/2\omega^2 ul) du.$$

Multiplying by $\exp izx$ and integrating over x first, we get for the characteristic function $\Phi(z)$

$$\Phi(z) = \{\lambda[\lambda + (\omega^2 z^2/2\omega_0^2)]^{-1}\}^{\nu}.$$

Taking the j th power of this function merely replaces ν by $j\nu$ and we see that the convolution is given by

$$P_j(x) = \frac{\lambda^{j\nu}}{\Gamma(j\nu)} \int_0^{\infty} u^{j\nu-1} \frac{\omega_0}{\omega\sqrt{2\pi lu}} \exp(-\lambda u - \omega_0^2 x^2/2\omega^2 ul) du. \quad (39)$$

For this distribution we have for the mean loss in db and the standard deviation in db

$$m_j = -(10 \log e) \left[\psi(j\nu) + \psi\left(\frac{1}{2}\right) \right] \quad (40)$$

$$+ 10 \log \lambda + 10 \log \frac{\omega_0^2}{2\omega^2} - 10 \log l$$

$$\sigma_j = (10 \log e) \left[\psi'(j\nu) + \psi'\left(\frac{1}{2}\right) \right]^{\frac{1}{2}}. \quad (41)$$

We may now apply these results to the general far-end crosstalk situation of Fig. 11. Suppose that the cable section in which the far-end

exposure takes place has loss L_1 and length L_1/K miles, where K is the loss in db per mile, and that it consists of j reels of length l miles, i.e., $jl = L_1/k$. The ratio of the power at the input of amplifier 2 to that arriving at the exposure section on the disturbing pair at frequency f , for a pair combination giving the mean dbm power, is

$$(2l/\lambda)(f/f_0)^2 \exp \left[\psi(j\nu) + \psi\left(\frac{1}{2}\right) - 2\alpha(f)(L_1/K) \right]$$

where $\alpha(f)$ is the cable loss in nepers per mile. Taking into account the preamplifier characteristic $r(f)$, the power spectrum $W(f)$ of the amplifier output signal as in (33), and the loss, L_2 , of the section of cable from amplifier 1 to the exposure, the total interference power at the preamplifier output for one disturber is

$$(2L_1/\lambda jK) \exp \left[\psi(j\nu) + \psi\left(\frac{1}{2}\right) \right] \int_0^\infty (f/f_0)^2 W(f) |r(f)|^2 \cdot \exp [-2\alpha(f)(L_1 + L_2)/K] df.$$

The interference power in dbm for this mean interferer is

$$\begin{aligned} \bar{Q}_1 = & 10 \log (2L_1/\lambda jK) + 10 \log e \left[\psi(j\nu) + \psi\left(\frac{1}{2}\right) \right] \\ & + 30 + 10 \log \int_0^\infty (f/f_0)^2 W(f) |r(f)|^2 \\ & \cdot \exp [-2\alpha(f)(L_1 + L_2)/K] df. \end{aligned} \quad (42)$$

The standard deviation of Q_1 is given by (41).

To determine the interfering power for n interferers, we can obtain from (39) the density for the square of the current and find its n -fold convolution. To obtain an approximate result we may use the Edson-Alford formula,²⁰ which gives an approximation to the distribution of the power sum expressed in db, of n quantities whose distribution in db is assumed to be Gaussian. The increase in the mean power in dbm for n interferers as compared to one interferer is

$$a_{j,n} = 5[\log (n^3 \exp C^2 \sigma_j^2) - \log (\exp C^2 \sigma_j^2 + n - 1)] \quad (43)$$

and the standard deviation is

$$\sigma_{j,n} = 6.593 [\log (\exp C^2 \sigma_j^2 + n - 1) - \log n]^{\frac{1}{2}} \quad (44)$$

where $C = 1/(10 \log e)$. Performing the numerical integration in (42) with the assumption that $\alpha(f) = \alpha(f_0)(f/f_0)^{\frac{1}{2}}$, we find that the integral

is very closely given by $7.0 - (L_1 + L_2)$. Thus the mean interfering power in dbm for n systems is

$$\bar{Q}_n = 10 \log (2L_1/\lambda jK) + 10 \log e [\psi(j\nu) + \psi(\frac{1}{2})] + 37.0 - (L_1 + L_2) + a_{j,n} \quad (45)$$

with standard deviation $\sigma_{j,n}$ as in (44).

The values of λ and ν in (45) must be determined by substituting in (40) and (41) the measured mean, m_{j_0} , and standard deviation σ_{j_0} , of the equal-level far-end crosstalk losses at a single frequency (f_0 is convenient) for a cable made up of j_0 reels of length l_0 miles each. If the loss m_{j_0} is measured at f_0 , we find that

$$10 \log \lambda = m_{j_0} + 10 \log e [\psi(j_0\nu) + \psi(\frac{1}{2})] + 10 \log 2l_0. \quad (46)$$

Substituting (46) in (45),

$$\bar{Q}_n = 10 \log (L_1/jKl_0) - m_{j_0} + 10 \log e [\psi(j\nu) - \psi(j_0\nu)] + 37.0 - (L_1 + L_2) + a_{j,n}. \quad (47)$$

Instead of L_1/jKl_0 , the equivalent expression l/l_0 may be used.

Equation (47) gives the mean interfering power, \bar{Q}_n , strictly speaking, only when the exposure section is made up of j reels, for $j = 1, 2, 3, \dots$. For exposure sections of arbitrary loss L_1 , and an assumed reel length of l miles, we may use $j = L_1/Kl$, which will not in general be an integer. The results for \bar{Q}_n and $\sigma_{j,n}$ will be interpolations between their values for the nearest integer j 's. For small values of L_1 when we would have $j < 1$, we cannot apply the convolution results (40) and (41); instead, the one-reel distribution applies. Thus, for $L_1/Kl < 1$, we have, putting $j = 1$ in (47),

$$\bar{Q}_n = 10 \log (L_1/Kl_0) - m_{j_0} + 10 \log e [\psi(\nu) - \psi(j_0\nu)] + 37.0 - (L_1 + L_2) + a_{1,n}. \quad (48)$$

The standard deviation of Q_n is $\sigma_{1,n}$.

To illustrate (47), suppose we have $n = 50$ systems in 22-gauge paper cable with $L = L_1 = 32.2$ db and $L_2 = 0$ in Fig. 11. If we take 900 feet as the reel length, the 32.2 db section will contain $j = 7$ reels. For the cable crosstalk parameters, we use the data of Table IV, choosing $m_{j_0} = 63$ db for pairs in the same unit. Other measurements²¹ on a single reel of cable gave a standard deviation σ_1 of 11 db. Substituting $\sigma_1 = 11$ in (41) shows that $\nu \approx 1$, and we adopt the value $\nu = 1$ for convenience ($\sigma_1 = 11.2$ db).

TABLE IV—FAR-END CROSSTALK DATA¹³

Type of cable:	22-gauge, paper-insulated, 900 pairs
Length:	6106 feet
Number of reels, j_0 :	9
Reel length, l_0 :	$6106/(9 \times 5280) = 0.13$ miles
Loss per mile, K :	26.5 db per mile (772 kc, 55 F)
Mean equal-level far-end crosstalk coupling loss, at 772 kc, m_{j_0} :	(a) for pairs in same unit, 63 db (b) for pairs in adjacent units, 77 db
Estimated parameter ν (see text)	1

From (41), we find that $\sigma_7 = 9.8$ db; putting this in (43), $a_{7,50} = 25.0$ db. The mean interference power \bar{Q}_{50} is then $= 33.2$ dbm, from (47). Further, its standard deviation $\sigma_{Q_{50}} = 5.2$ dbm, using (44).

APPENDIX C

Central Office Noise

The operation of electromagnetic relays in telephone offices produces electrical transients which propagate down the pairs.²² Through crosstalk in the cable, these disturbances may appear on pairs not connected to switches. The energy distribution of this noise is very broad, reaching frequencies well in excess of two megacycles. The transients are of several types and are quite complex. If T1 lines were operated at a lower signal level they would cause bursts of errors, each burst lasting less than a millisecond and containing from one or two to several hundred errors. Since the noise is primarily due to switching, the error rate is strongly dependent upon office activity, being highest during the busy hours and disappearing in the early hours of the morning. Since a crosstalk path is involved, the levels of the noise transients reflect the approximately log-normal distribution of crosstalk loss and vary widely.

Because of the complexity of the office noise phenomenon, it is necessary to lean heavily on experimental results. The feature of the noise which is of most interest in T1 carrier work is the error rate which it produces in the office amplifier, as a function of the length of the repeater section. The error rate is predictable from the amplitude distribution of the noise at the output of the linear preamplifier, that is, at the point in the amplifier at which the decision is made whether a pulse has been received or not. The amplifier may be assumed to make an error if the noise amplitude is greater than one-fourth the pulse height (Section 2.1). The error rate to be expected on a typical office repeater section has therefore been studied by means of a sampling and comparison device which is

connected to an otherwise idle pair at the main frame. A noise sample increases the count on a register by one if the amplitude is greater than a preset level.

Since the repeater section is always built out to full length (Section 2.2), the signal pulse shape and amplitude at the preamplifier output are always the same, within the allowable ± 4 db limits. The noise, on the other hand, is attenuated by whatever buildout is used. Since the buildout loss is a function of frequency and the noise transients are quite variable in structure, the relation between section loss and error rate must also be found experimentally. The noise may also be attenuated by pair loss* in the disturbing pair if the amplifier is not in the office but near it. Since the buildouts are designed to imitate pair loss, the effect on the error rate of adding pair loss in the noise path is the same as the effect of shortening the repeater section.

In some cases the noise exposure occurs via far-end rather than near-end crosstalk. The dependence on the crosstalk mode may be estimated from the relative near-end and far-end crosstalk losses. Again, if the disturbing pairs are in a different unit or layer of the cable than the T1 pairs, the loss in the crosstalk path is greater than if both are in the same unit, and this may be taken into account in engineering end sections.

Since both crosstalk interference and impulse noise are generally present in sections near offices, the error rate due to any combination of the two must be determined.

To summarize, the following questions must be answered in order to estimate the restrictions which should be placed on repeater sections exposed to office noise:

- (a) What is the worst noise level to be expected in cable pairs entering an office?
- (b) How does noise error rate depend upon section loss, distance of the repeater from the office, crosstalk mode, and segregation in the sheath?
- (c) If office noise and crosstalk separately produce error rates R_1 and R_2 , what error rate will both together produce?

Considering the first of these questions, amplitude distributions for positive noise voltages were measured *at the output of a T1 preamplifier* on several pairs at their main frame terminations in one office. Curve 1 of Fig. 19 is for the median pair. The ordinate is the normal probability scale, while the abscissa gives the instantaneous level at the output of the preamplifier in dbm, with no buildout in the amplifier. Thus a level dis-

* Noise in the disturbing pair is found in both the longitudinal and metallic modes. At T1 frequencies the losses in the two modes may be taken to be the same for the present purpose.

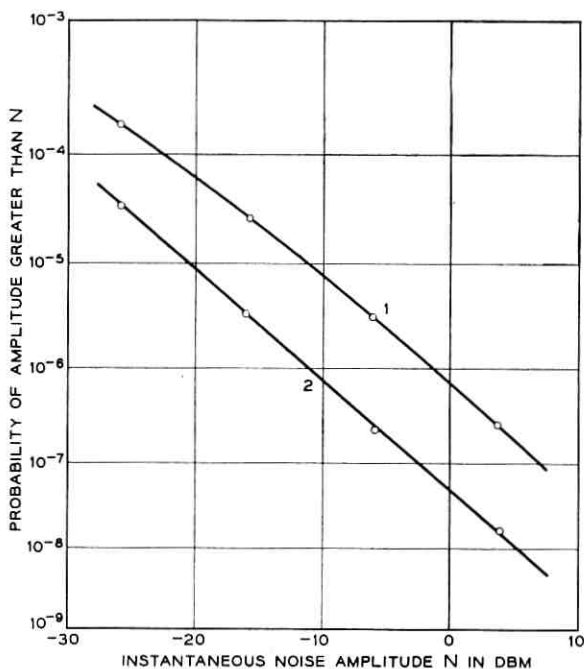


Fig. 19 — Noise amplitude distribution at preamplifier output with 9.6-db buildout (curve 2) and without (curve 1).

tribution which was normal in db would appear as a straight line. Of several offices involving step-by-step, panel, and crossbar switching machines, this office gave as high a noise level as any. We have, therefore, taken curve 1 of Fig. 19 to be representative of noise levels in noisy offices.

Considering the second question, if section loss is decreased, the noise level is reduced by the presence of the buildout. Curve 2 of Fig. 19 shows the effect of the 9.6 db buildout on the noise distribution of curve 1. The effect is quite closely the same as the introduction of the same amount of flat loss. Thus reducing the section loss or moving the repeater away from the office affects the noise level by the number of db represented by the change, measured at 772 kc.

For calculations based on Fig. 19, we assume that the mean near-end coupling loss is 75 db at 772 kc, a figure that is typical for pairs in the same unit or layer of large cables.¹³ In the case of far-end coupling, the approximation

$$\bar{L}_f \approx 63 - 10 \log (l/6106) \quad (49)$$

may be used; \bar{L}_f is the mean equal-level far-end crosstalk coupling loss at 772 kc for a cable of length l feet (cf. Appendix B). To this loss is added the direct loss of the pair between the office and the disturbed amplifier.

The additional loss in the crosstalk path due to segregation of T1 pairs in a unit or layer of the cable containing no switched pairs can be estimated from the relative mean crosstalk losses. Near-end crosstalk loss in unit cables is about 8 to 13 db greater in mean value for pairs in adjacent units than for pairs in the same unit. Far-end crosstalk loss is about 10 to 15 db greater for pairs in adjacent units than for pairs in the same unit. These are metallic-to-metallic losses, whereas longitudinal-to-metallic couplings may be dominant. Some noise data is available which indicates about a 10-db reduction of office noise for pairs in an idle unit in the sheath as compared with pairs in working units. This estimate has been used for both near-end and far-end noise coupling.

Curve 1 of Fig. 19 may now be translated into a curve of error rate as a function of the difference between loss in the noise path and loss in the signal path. For example, the curve shows that a noise level of -1.5 dbm at the preamplifier output is exceeded with probability 10^{-6} . Therefore, the amplifier error rate will be 10^{-6} when the signal-to-noise ratio at this point is 12 db (Section 2.1), i.e. when the signal level is $-1.5 + 12 = 10.5$ dbm. The signal level (Section 2.2) at the preamplifier output is $\bar{S} + G - \bar{L}$, where these symbols have the same meaning as in Section 4.2.1. Substituting $\bar{S} = 18.6$ dbm and $G = 23.7$, we find that the mean pair loss, i.e. the loss in the signal path, is $\bar{L} = 31.8$ db for a signal level of 10.5 dbm. We have assumed a mean loss in the noise path of 75 db in Fig. 19. The difference between noise path loss and signal path loss is thus $75 - 31.8 = 43.2$ db. In Fig. 19, error rate 10^{-6} is plotted above the abscissa 43.2 db. Correspondingly, each abscissa of Fig. 19 is increased by $43.2 + 1.5 = 44.7$ db to obtain a point of Fig. 20. When the losses in the noise and signal paths for a particular amplifier are known, the error rate in the presence of severe office noise may be read from Fig. 20.

Finally, the error rate resulting from a combination of crosstalk and noise must be examined. This requires choosing a mathematical representation of the office noise amplitude probability density p_n and finding its convolution with the normal distribution, which applies to crosstalk voltage amplitudes.

A convenient representation for this purpose is:

$$p_n(x) = \frac{\sqrt{2}\alpha^3}{\pi} (\alpha^4 + x^4)^{-1}. \quad (50)$$

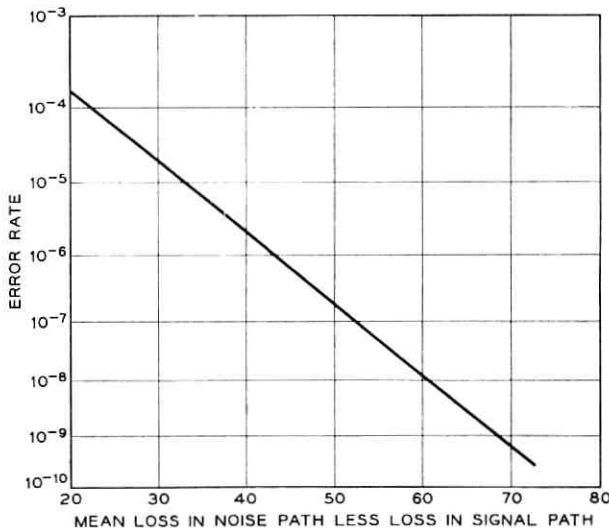


Fig. 20 — Error rate as a function of loss difference.

This function fits typical noise curves quite well. The fit is not as good for the rather extreme case of curve 1 of Fig. 19, which shows a relatively higher probability of very large amplitudes, but it is satisfactory for the present purpose. For large values of the ratio of the random amplitude variable X to the standard deviation α we have

$$\text{Prob} \left(\frac{X}{\alpha} \geq u \right) = \frac{\sqrt{2}}{\pi} \left\{ \frac{1}{3u^3} - \frac{1}{7u^7} + \frac{1}{11u^{11}} + \dots \right\}.$$

For an error rate of 10^{-7} all terms except the first can be ignored and we find $u = 114.5$. The probability density of the amplitude of the crosstalk interference at the preamplifier output is given by

$$p_c(x) = \frac{1}{\sqrt{2\pi}\sigma} e^{-x^2/2\sigma^2}. \tag{51}$$

If Y is the crosstalk interference amplitude variable, solving the equation

$$\text{Prob} \left(\frac{Y}{\sigma} \geq u \right) = 10^{-7}$$

yields $u = 5.2$. Thus an error rate of 10^{-7} will occur if the error threshold (one-fourth the pulse height) is 5.2 times the rms crosstalk interference or if it is 114.5 times the rms office noise. We are interested in the error

rate if these amounts of noise and crosstalk interference occur in the same amplifier. This may be calculated by finding the distribution of the sum $Y + X$ of the two variables. Let

$$\alpha = \frac{5.2}{114.5} \sigma.$$

By working with the characteristic functions corresponding to the probability densities of (50) and (51) it may be shown that

$$\text{Prob}(X + Y \geq 5.2\sigma) = 2.31 \times 10^{-7}.$$

Thus the error rate due to the sum of the two types of interference is approximately the sum of the error rates due to the two interferences separately. For practical purposes, design requirements for office noise and crosstalk may therefore be considered separately.

REFERENCES

1. Davis, C. G., B.S.T.J., **41**, 1962, pp. 1-24.
2. Fultz, K. E., and Williams, O. L., B.S.T.J., to appear.
3. Mayo, J. S., B.S.T.J., **41**, 1962, pp. 25-97.
4. Aaron, M. R., B.S.T.J., **41**, 1962, pp. 99-141.
5. Caruthers, R. S., B.S.T.J., **30**, 1951, pp. 1-32.
6. Edwards, P. G., and Montfort, L. R., B.S.T.J., **31**, 1952, pp. 688-723.
7. Oliver, B. M., Pierce, J. R., and Shannon, C. E., Proc. IRE, **36**, 1948, pp. 1324-1331.
8. Hölzler, E., and Holzwarth, H., *Theorie und Technik der Pulsmodulation*, Springer, Berlin, 1957.
9. Eager, G. S., Jr., Jachimowicz, L., Kolodny, I., and Robinson, D. E., Trans. AIEE, pt. I, **78**, 1959, pp. 618-640.
10. Piper, H. S. Jr., unpublished.
11. Henneberger, T. C., and Fagen, M. D., Commun. and Electronics, No. 59, 1962, pp. 27-33.
12. Valley, G. E., Jr., and Wallman, H., eds., *Vacuum Tube Amplifiers*, McGraw-Hill, New York, 1948.
13. Mallett, J., unpublished.
14. Robinson, D. B., Jr., unpublished.
15. DeWitt, R. G., unpublished.
16. Campbell, G. A., B.S.T.J., **14**, 1935, pp. 558-572.
17. Kaden, H., Archiv der Elektr. Übertragung, **11**, 1957, pp. 349-354.
18. *Jahnke-Emde-Lösch Tables of Higher Functions*, sixth ed., McGraw-Hill, New York, 1960.
19. Davis, H. T., *Tables of the Higher Mathematical Functions*, Principia, Bloomington, Indiana, 1933, 1935.
20. Holbrook, B. D., unpublished.
21. Rentrop, E., unpublished.
22. Curtis, A. M., B.S.T.J., **19**, 1940, pp. 40-62.

Delay Distributions for One Line with Poisson Input, General Holding Times, and Various Orders of Service

By L. TAKÁCS

(Manuscript received December 14, 1962)

At a telephone exchange, calls appear before a single trunk line in accordance with a Poisson process of density λ . If the trunk line is busy, calls are delayed. The call holding times are identically distributed, mutually independent, positive random variables with distribution function $H(x)$. In this paper the distribution function of the delay and its moments are given for a stationary process and for three orders of service: (i) order of arrival, (ii) random order, and (iii) reverse order of arrival.

I. INTRODUCTION

Let us suppose that in the time interval $(0, \infty)$ calls appear before a single trunk line at times $\tau_1, \tau_2, \dots, \tau_n, \dots$ where the interarrival times $\tau_n - \tau_{n-1}$ ($n = 1, 2, \dots; \tau_0 = 0$) are identically distributed, mutually independent random variables with the distribution function

$$F(x) = \begin{cases} 1 - e^{-\lambda x} & \text{if } x \geq 0, \\ 0 & \text{if } x < 0, \end{cases} \quad (1)$$

that is, the input is a Poisson process of density λ . If an incoming call finds the line free, a connection is realized instantaneously. If the line is busy, the call is delayed and waits for service as long as necessary (no defection). The holding times are identically distributed, mutually independent, positive random variables with distribution function $H(x)$ and independent of the input process. Such a service system can be characterized by the symbol $[F(x), H(x), 1]$ provided that the order of service is specified. In this paper three orders of service are considered: (i) *order of arrival* (first come-first served), (ii) *random order* (every waiting call, independently of the others, and of its past delay, has the same probability of being chosen for service), and (iii) *reverse order of arrival* (last come-first served).

We are interested in finding the distribution function of the delay for a stationary process and for the three orders of service. We shall prove that if $\lambda\alpha < 1$, where α is the average holding time, then there is a unique stationary process.

Throughout this paper we shall use the notation

$$\Psi(s) = \int_0^{\infty} e^{-sx} dH(x) \quad (\Re(s) \geq 0) \quad (2)$$

and

$$\alpha_k = \int_0^{\infty} x^k dH(x) \quad (k = 0, 1, 2, \dots). \quad (3)$$

In particular, $\alpha_1 = \alpha$ is the average holding time.

II. THE STATIONARY PROCESS

Let us denote by ξ_n the queue size at time $t = \tau_n - 0$, i.e., the n th incoming call finds ξ_n calls (either waiting or being served) in the system. Denote by χ_n the time needed to complete the current service (if any) at time $t = \tau_n - 0$. If $\xi_n = 0$ then $\chi_n = 0$. The vector sequence

$$(\xi_n, \chi_n), \quad n = 1, 2, \dots,$$

is a Markovian stochastic sequence and has the same stochastic behavior for each order of service. We shall prove that if $\lambda\alpha < 1$ then there exists a unique stationary distribution, whereas if $\lambda\alpha \geq 1$ then a stationary distribution does not exist. If $\lambda\alpha < 1$ and (ξ_1, χ_1) has the stationary distribution, then every (ξ_n, χ_n) has the same distribution as the initial distribution. For the stationary process, let us introduce the following notation

$$\mathbf{P}\{\xi_n = j\} = P_j \quad (j = 0, 1, \dots) \quad (4)$$

$$\mathbf{P}\{\chi_n \leq x, \xi_n = j\} = P_j(x) \quad (x \geq 0, j = 1, 2, \dots) \quad (5)$$

and

$$\Pi_j(s) = \int_0^{\infty} e^{-sx} dP_j(x) \quad (\Re(s) \geq 0, j = 1, 2, \dots). \quad (6)$$

We shall prove the following theorem, due to D. M. G. Wishart:¹

Theorem 1: If $\lambda\alpha < 1$, then the stochastic sequence (ξ_n, χ_n) , $n = 1, 2, \dots$, has a unique stationary distribution which is given by $P_0 = 1 - \lambda\alpha$ and

$$\begin{aligned}
 U(s, z) &= \sum_{j=1}^{\infty} \Pi_j(s) z^j \\
 &= \frac{(1 - \lambda\alpha)\lambda z(1 - z)}{z - \Psi(\lambda(1 - z))} \left(\frac{\Psi(s) - \Psi(\lambda(1 - z))}{s - \lambda(1 - z)} \right) \tag{7}
 \end{aligned}$$

for $\Re(s) \geq 0$ and $|z| \leq 1$.

Proof: If we express the distribution of (ξ_{n+1}, χ_{n+1}) with the aid of the distribution of (ξ_n, χ_n) , and assume that both (ξ_{n+1}, χ_{n+1}) and (ξ_n, χ_n) have the same stationary distribution, and if we form Laplace-Stieltjes transforms, then we obtain that P_0 and $\Pi_j(s)$ ($j = 1, 2, \dots$) must satisfy the following system of linear equations:

$$\begin{aligned}
 P_0 &= P_0\Psi(\lambda) + \sum_{k=1}^{\infty} \Pi_k(\lambda)[\Psi(\lambda)]^k, \\
 \Pi_1(s) &= \frac{\lambda[\Psi(\lambda) - \Psi(s)]}{s - \lambda} \left\{ P_0 + \sum_{k=1}^{\infty} \Pi_k(s)[\Psi(\lambda)]^{k-1} \right\}, \\
 \Pi_j(s) &= \frac{\lambda[\Pi_{j-1}(\lambda) - \Pi_{j-1}(s)]}{s - \lambda} \\
 &\quad + \frac{\lambda[\Psi(\lambda) - \Psi(s)]}{s - \lambda} \sum_{k=j}^{\infty} \Pi_k(s)[\Psi(\lambda)]^{k-j}
 \end{aligned} \tag{8}$$

for $j = 2, 3, \dots$ and

$$P_0 + \sum_{j=1}^{\infty} \Pi_j(0) = 1. \tag{9}$$

To prove (8) we use the following two facts: First, the probability that during a holding time no call arrives is given by

$$\Psi(\lambda) = \int_0^{\infty} e^{-\lambda x} dH(x).$$

Second, let ρ and θ be mutually independent, positive random variables with distribution functions $\mathbf{P}\{\rho \leq x\} = P(x)$, and $\mathbf{P}\{\theta \leq x\} = F(x)$ defined by (1). Write $A = \{\theta < \rho\}$. Then

$$\mathbf{P}\{A\} \mathbf{E}\{e^{-s(\rho-\theta)} | A\} = \int_0^{\infty} \int_0^x e^{-(x-y)s-\lambda y} dy dP(x) = \frac{\lambda[\Pi(\lambda) - \Pi(s)]}{s - \lambda}$$

where

$$\Pi(s) = \int_0^{\infty} e^{-sx} dP(x).$$

Forming generating functions in (8) we obtain

$$[s - \lambda(1 - z)]U(s, z) = \lambda z U(\lambda, z) + \lambda z [\Psi(\lambda) - \Psi(s)] \left\{ P_0 + \frac{U(\lambda, \Psi(\lambda)) - U(\lambda, z)}{\Psi(\lambda) - z} \right\}. \quad (10)$$

If $s = \lambda(1 - z)$ in (10) then we get

$$U(\lambda, z) + [\Psi(\lambda) - \Psi(\lambda(1 - z))] \left\{ P_0 + \frac{U(\lambda, \Psi(\lambda)) - U(\lambda, z)}{\Psi(\lambda) - z} \right\} = 0. \quad (11)$$

The comparison of (10) and (11) gives

$$U(s, z) = \frac{\lambda z [\Psi(s) - \Psi(\lambda(1 - z))]}{[s - \lambda(1 - z)][\Psi(\lambda) - \Psi(\lambda(1 - z))]} U(\lambda, z). \quad (12)$$

By the first equation of (8), $U(\lambda, \Psi(\lambda)) = P_0[1 - \Psi(\lambda)]$, and if we put this into (11) we get

$$U(\lambda, z) = \frac{P_0(1 - z)[\Psi(\lambda) - \Psi(\lambda(1 - z))]}{z - \Psi(\lambda(1 - z))}. \quad (13)$$

Thus by (12) and (13)

$$U(s, z) = \frac{\lambda P_0 z (1 - z) [\Psi(s) - \Psi(\lambda(1 - z))]}{[z - \Psi(\lambda(1 - z))][s - \lambda(1 - z)]}. \quad (14)$$

Since by (9) $P_0 + U(0, 1) = 1$, it follows from (14) that $P_0 = 1 - \lambda\alpha$. Thus if $\lambda\alpha \geq 1$, then the assumption that a stationary distribution exists leads to a contradiction, i.e., a stationary distribution cannot exist if $\lambda\alpha \geq 1$. If $\lambda\alpha < 1$, then there exists one and only one stationary distribution which is given by $P_0 = 1 - \lambda\alpha$ and by (14). This proves (7).

Remark. From (7) we obtain by inversion that for $x \geq 0$

$$\begin{aligned} \sum_{j=1}^{\infty} P_j(x) z^j &= \frac{(1 - \lambda\alpha)\lambda z(1 - z)}{z - \Psi(\lambda(1 - z))} \int_0^{\infty} e^{-\lambda(1-z)u} [H(u + x) - H(u)] du. \end{aligned} \quad (15)$$

Hence for $x \geq 0$

$$\sum_{j=1}^{\infty} P_j(x) = \lambda \int_0^x [1 - H(u)] du. \quad (16)$$

Accordingly if (ξ_n, χ_n) , $n = 1, 2, \dots$, is a stationary sequence, then we have for $x \geq 0$ that

$$\mathbf{P}\{\chi_n \leq x \mid \xi_n \geq 1\} = \frac{1}{\alpha} \int_0^x [1 - H(u)] du, \tag{17}$$

i.e., if an incoming call finds the line busy, then the distribution function of the time needed to complete the current service is given by

$$H^*(x) = \begin{cases} \frac{1}{\alpha} \int_0^x [1 - H(u)] du & \text{if } x \geq 0, \\ 0 & \text{if } x < 0. \end{cases} \tag{18}$$

Finally we also remark that the stationary distribution of ξ_n , $n = 1, 2, \dots$, is given by the following generating function

$$U(z) = \sum_{j=0}^{\infty} P_j z^j = \frac{(1 - \lambda\alpha)(1 - z)\Psi(\lambda(1 - z))}{\Psi(\lambda(1 - z)) - z}. \tag{19}$$

This follows from (7), because $U(z) = P_0 + U(0, z)$.

III. THE DISTRIBUTION FUNCTION OF THE LENGTH OF A BUSY PERIOD

A busy period is defined as a time interval during which the line is continuously busy. The stochastic law of a busy period is obviously independent of the order of service. Every busy period (except the initial one, if the line is busy at time $t = 0$) independently of the others has the same stochastic law. Denote by $G(x)$ the probability that the length of a busy period (other than the initial one, if any) is $\leq x$ and define

$$\gamma(s) = \int_0^{\infty} e^{-sx} dG(x) \quad (\Re(s) \geq 0). \tag{20}$$

In Ref. 2 it is proved that $\gamma(s)$ is the root with smallest absolute value in z of the equation

$$z = \Psi(s + \lambda(1 - z)). \tag{21}$$

By Lagrange's expansion (cf. Ref. 3, p. 132) we obtain that

$$\gamma(s) = \sum_{n=1}^{\infty} \frac{\lambda^{n-1}}{n!} \int_0^{\infty} e^{-(\lambda+s)x} x^{n-1} dH_n(x) \tag{22}$$

where $H_n(x)$ denotes the n th iterated convolution of $H(x)$ with itself. From (22) it follows by inversion that

$$G(x) = \sum_{n=1}^{\infty} \frac{\lambda^{n-1}}{n!} \int_0^x e^{-\lambda u} u^{n-1} dH_n(u). \tag{23}$$

If $\lambda\alpha \leq 1$ then $G(\infty) = 1$, whereas if $\lambda\alpha > 1$ then $G(\infty) < 1$.

In the case of $\lambda\alpha \leq 1$ the r th moment of $G(x)$ is defined by

$$\Gamma_r = \int_0^\infty x^r dG(x) \quad (r = 0, 1, 2, \dots). \quad (24)$$

If $\lambda\alpha < 1$ and α_r is finite then $\Gamma_0, \Gamma_1, \dots, \Gamma_r$ are also finite and we have $\Gamma_0 = 1, \Gamma_1 = \alpha/(1 - \lambda\alpha)$, and

$$\Gamma_{n+1} = \sum_{\nu=1}^n \frac{(n + \nu)! \lambda^{\nu-1}}{n! (1 - \lambda\alpha)^{n+\nu+1}} Y_{n,\nu} \quad (25)$$

for $n = 1, 2, \dots$, where

$$Y_{n,\nu} = \sum_{\substack{j_1+j_2+\dots+j_n=\nu \\ j_1+2j_2+\dots+nj_n=n}} \frac{n! \alpha_2^{j_1} \alpha_3^{j_2} \dots \alpha_{n+1}^{j_n}}{j_1! j_2! \dots j_n! (2!)^{j_1} (3!)^{j_2} \dots ((n+1)!)^{j_n}}. \quad (26)$$

If, in particular, $H(x) = 1 - e^{-x/\alpha}$ ($x \geq 0$), then $\alpha_r = r! \alpha^r$ and

$$Y_{n,\nu} = \alpha^{n+\nu} \sum_{\substack{j_1+j_2+\dots+j_n=\nu \\ j_1+2j_2+\dots+nj_n=n}} \frac{n!}{j_1! j_2! \dots j_n!} = \frac{n!}{\nu!} \binom{n-1}{\nu-1} \alpha^{n+\nu}.$$

Formula (26) can be proved as follows. If we define

$$u = s + \lambda[1 - \gamma(s)],$$

then by (21) $s = u - \lambda[1 - \Psi(u)]$, whence by Bürmann's theorem (cf. Appendix) for $n = 0, 1, \dots$ we have

$$\left(\frac{d^{n+1}u}{ds^{n+1}} \right)_{s=0} = \left[\frac{d^n}{du^n} \left(\frac{u}{s} \right)^{n+1} \right]_{u=0} = \left[\frac{d^n}{du^n} \left(\frac{1}{1 - \lambda \frac{1 - \Psi(u)}{u}} \right)^{n+1} \right]_{u=0} \quad (27)$$

and the n th derivative can be calculated by using Faa di Bruno's formula (cf. Appendix). On the other hand

$$\left(\frac{du}{ds} \right)_{s=0} = 1 - \lambda\gamma'(0) = 1 + \lambda\Gamma_1, \quad (28)$$

and

$$\left(\frac{d^{n+1}u}{ds^{n+1}} \right)_{s=0} = -\lambda\gamma^{(n+1)}(0) = (-1)^n \lambda \Gamma_{n+1} \quad (n = 1, 2, \dots). \quad (29)$$

Comparing the above formulas we obtain Γ_n for every n .

Finally we remark that, by (25)

$$\Gamma_2 = \frac{\alpha_2}{(1 - \lambda\alpha)^3}, \quad (30)$$

$$\Gamma_3 = \frac{\alpha_3}{(1 - \lambda\alpha)^4} + \frac{3\lambda\alpha_2^2}{(1 - \lambda\alpha)^5}, \tag{31}$$

$$\Gamma_4 = \frac{\alpha_4}{(1 - \lambda\alpha)^5} + \frac{10\lambda\alpha_2\alpha_3}{(1 - \lambda\alpha)^6} + \frac{15\lambda^2\alpha_2^3}{(1 - \lambda\alpha)^7}. \tag{32}$$

IV. THE DISTRIBUTION FUNCTION OF THE DELAY

Let us denote by η_n the delay of the n th call. If the order of service is specified then the distribution function of η_n is uniquely determined by the distribution of (ξ_n, χ_n) . If (ξ_n, χ_n) , $n = 1, 2, \dots$, is a stationary stochastic sequence, then η_n has the same distribution for every n . In the case of the stationary process write $\mathbf{P}\{\eta_n \leq x\} = W(x)$ and

$$\mathbf{E}\{e^{-s\eta_n}\} = \Omega(s)$$

for each order of service. Define

$$W_n = \int_0^\infty x^n dW(x) \quad (n = 0, 1, 2, \dots). \tag{33}$$

In each case W_n is finite if α_{n+1} is finite. For each order of service

$$W_1 = \frac{\lambda\alpha_2}{2(1 - \lambda\alpha)}. \tag{34}$$

For service in order of arrival

$$W_2 = \frac{\lambda\alpha_3}{3(1 - \lambda\alpha)} + \frac{\lambda^2\alpha_2^2}{2(1 - \lambda\alpha)^2}, \tag{35}$$

for service in random order

$$W_2 = \frac{2\lambda\alpha_3}{3(1 - \lambda\alpha)(2 - \lambda\alpha)} + \frac{\lambda^2\alpha_2^2}{(1 - \lambda\alpha)^2(2 - \lambda\alpha)}, \tag{36}$$

and for service in reverse order of arrival

$$W_2 = \frac{\lambda\alpha_3}{3(1 - \lambda\alpha)^2} + \frac{\lambda^2\alpha_2^2}{2(1 - \lambda\alpha)^3}. \tag{37}$$

(i) *Service in Order of Arrival.* This case was first investigated by F. Pollaczek⁴ and A. Y. Khintchine.⁵ Cf. also D. V. Lindley.⁶

Theorem 2: If $\lambda\alpha < 1$, if the process is stationary and if service is in order of arrival, then the distribution function of the delay of a call is given by

$$W(x) = (1 - \lambda\alpha) \sum_{k=0}^\infty (\lambda\alpha)^k H_k^*(x) \tag{38}$$

where $H_k^*(x)$ denotes the k -th iterated convolution of

$$H^*(x) = \begin{cases} \frac{1}{\alpha} \int_0^x [1 - H(u)] du & \text{if } x \geq 0, \\ 0 & \text{if } x < 0, \end{cases} \quad (39)$$

with itself; $H_0^*(x) = 1$ if $x \geq 0$ and $H_0^*(x) = 0$ if $x < 0$.

Proof: Evidently

$$W(x) = P_0 + \sum_{j=1}^{\infty} P_j(x) * H_{j-1}(x)$$

where $H_j(x)$ ($j = 1, 2, \dots$) denotes the j th iterated convolution of $H(x)$ with itself; $H_0(x) = 1$ if $x \geq 0$ and $H_0(x) = 0$ if $x < 0$. The symbol $*$ denotes convolution. Hence

$$\Omega(s) = P_0 + \sum_{j=1}^{\infty} \Pi_j(s) [\Psi(s)]^{j-1} = P_0 + \frac{U(s, \Psi(s))}{\Psi(s)} \quad (41)$$

where $U(s, z)$ is defined by (7) and $P_0 = 1 - \lambda\alpha$. Thus

$$\Omega(s) = \frac{1 - \lambda\alpha}{1 - \lambda \frac{1 - \Psi(s)}{s}}, \quad (42)$$

whence (38) follows by inversion. Formula (38) was found by V. E. Beneš.⁷

If α_{n+1} is finite then W_n is also finite and is given by

$$W_n = \sum_{\nu=1}^n \frac{\lambda^\nu \nu!}{(1 - \lambda\alpha)^\nu} Y_{n,\nu} \quad (43)$$

where $Y_{n,\nu}$ is defined by (26). For,

$$W_n = (-1)^n \left(\frac{d^n \Omega(s)}{ds^n} \right)_{s=0} \quad (n = 0, 1, \dots) \quad (44)$$

and the n th derivative of $\Omega(s)$ can be obtained by Faa di Bruno's formula (cf. Appendix).

In this case W_1 is given by (34), W_2 by (35), and

$$W_3 = \frac{\lambda\alpha_4}{4(1 - \lambda\alpha)} + \frac{\lambda^2 \alpha_2 \alpha_3}{(1 - \lambda\alpha)^2} + \frac{3\lambda^3 \alpha_2^3}{4(1 - \lambda\alpha)^3}. \quad (45)$$

Remark. Let $T(x) = W(x) * H(x)$, i.e., $T(x)$ is the distribution function of the sum of the delay and the holding time of a call for a stationary process. Define

$$T_n = \int_0^\infty x^n dT(x) \quad (n = 0, 1, 2, \dots). \quad (46)$$

If α_{n+1} is finite then T_n is also finite and is given by $T_1 = W_1 + \alpha$ and

$$T_n = W_n + \frac{n}{\lambda} W_{n-1} \quad (n = 2, 3, \dots) \quad (47)$$

where W_n is defined by (43). Conversely, if we know T_j for $j = 1, 2, \dots, n$ then we can obtain W_n by the following formula

$$W_n = \frac{(-1)^n n!}{\lambda^n} \left[\lambda\alpha + \sum_{j=1}^n \frac{(-1)^j \lambda^j T_j}{j!} \right] \quad (n = 1, 2, \dots). \quad (48)$$

Formulas (47) and (48) follow from the relationship

$$\int_0^\infty e^{-sx} dT(x) = \Omega(s)\Psi(s) = \frac{(1 - \lambda\alpha)s}{\lambda} + \left(1 - \frac{s}{\lambda}\right)\Omega(s). \quad (49)$$

Finally we also note that the r th binomial moment of the stationary distribution of the queue size, i.e., that of $\{P_j\}$ defined by (19), is given by

$$B_r = \sum_{j=r}^\infty \binom{j}{r} P_j = \frac{\lambda^r T_r}{r!} \quad (r = 0, 1, \dots). \quad (50)$$

For we can easily see that

$$P_j = \int_0^\infty e^{-\lambda x} \frac{(\lambda x)^j}{j!} dT(x), \quad (51)$$

whence

$$B_r = \int_0^\infty \frac{(\lambda x)^r}{r!} dT(x) = \frac{\lambda^r T_r}{r!}. \quad (52)$$

(ii) *Service in Random Order.* The case of exponentially distributed holding times was investigated by many authors (cf. Ref. 8), the case of constant holding time by P. J. Burke,⁹ and the general case by J. F. Kingman.¹⁰ The following theorem is due to J. F. Kingman.¹⁰

Theorem 3: If $\lambda\alpha < 1$, if the process is stationary and if service is in random order, then the distribution function of the delay of a call has the following Laplace-Stieltjes transform:

$$\Omega(s) = (1 - \lambda\alpha) \left\{ 1 + \frac{\lambda}{s} \int_{\gamma(s)}^1 \exp \left[- \int_u^1 \frac{dv}{v - \Psi(s + \lambda(1 - v))} \right] \cdot \left[1 + \frac{u - 1}{u - \Psi(\lambda(1 - u))} \right] du \right\} \quad (53)$$

where $\gamma(s)$ is the root with smallest absolute value in z of the equation

$$z = \Psi(s + \lambda(1 - z)) \quad (54)$$

and is given by (22).

Proof: Under the condition that j ($j = 1, 2, \dots$) calls are waiting in the system when a service is about to start, denote by $W_j(x)$ the probability that the service of a given call among the j calls starts within time x if time is measured from this instant. Define

$$\Omega_j(s) = \int_0^\infty e^{-sx} dW_j(x) \quad (\Re(s) \geq 0). \quad (55)$$

The distribution functions $W_j(x)$ ($j = 1, 2, \dots$) can be obtained by using the following relationships: $W_1(x) = 1$ and

$$W_j(x) = \frac{1}{j} + \left(1 - \frac{1}{j}\right) \sum_{k=0}^{\infty} \left[\int_0^x e^{-\lambda u} \frac{(\lambda u)^k}{k!} dH(u) \right] * W_{j+k-1}(x) \quad (56)$$

for $j = 2, 3, \dots$ and for $x \geq 0$. To prove (56) we take into consideration that if the given call will be chosen for service among the j waiting calls, then its service starts immediately; if the given call is not chosen for service at this time then it must wait during the holding time of the call chosen for service, and if during this holding time k new calls arrive, then there is an additional delay which has the distribution function $W_{j+k-1}(x)$. Forming the Laplace-Stieltjes transform of (56) we obtain the following system of linear equations for the determination of $\Omega_j(s)$ ($j = 1, 2, \dots$): $\Omega_1(s) = 1$ and

$$j\Omega_j(s) = 1 + (j-1) \sum_{k=0}^{\infty} \Omega_{j+k-1}(s) \int_0^\infty e^{-(\lambda+s)x} \frac{(\lambda x)^k}{k!} dH(x) \quad (57)$$

for $j = 2, 3, \dots$. The solution of this system is given by J. F. Kingman¹⁰ in the following form:

$$\Omega_j(s) = \int_{\gamma(s)}^1 \exp \left[- \int_u^1 \frac{dv}{v - \Psi(s + \lambda(1 - v))} \right] \cdot \frac{u^{j-1}}{u - \Psi(s + \lambda(1 - u))} du \quad (58)$$

where $\gamma(s)$ is defined by (22). By integrating by parts (58) can be written in the following equivalent form:

$$\Omega_j(s) = 1 - (j-1) \int_{\gamma(s)}^1 \exp \left[- \int_u^1 \frac{dv}{v - \Psi(s + \lambda(1 - v))} \right] u^{j-2} du. \quad (59)$$

Now the distribution function of the delay is given by

$$W(x) = P_0 + \sum_{j=1}^{\infty} \sum_{k=0}^{\infty} \left[\int_0^x e^{-\lambda u} \frac{(\lambda u)^k}{k!} dP_j(u) \right] * W_{j+k}(x). \quad (60)$$

For, if a call arrives and finds the line free, then its service starts without delay. If an arriving call finds j ($j = 1, 2, \dots$) calls in the system, then its delay is composed of the time needed to complete the current service, and if during this time k new calls arrive, then there is an additional delay that has the distribution function $W_{j+k}(x)$. Forming the Laplace-Stieltjes transform of (60) we get

$$\Omega(s) = P_0 + \sum_{j=1}^{\infty} \sum_{k=0}^{\infty} \left[\int_0^{\infty} e^{-(\lambda+s)x} \frac{(\lambda x)^k}{k!} dP_j(x) \right] \Omega_{j+k}(s). \quad (61)$$

Putting (58) into (61), we obtain

$$\Omega(s) = P_0 + \int_{\gamma(s)}^1 \exp \left[- \int_u^1 \frac{dv}{v - \Psi(s + \lambda(1 - v))} \right] \cdot \frac{U(s + \lambda(1 - u), u)}{u[u - \Psi(s + \lambda(1 - u))]} du \quad (62)$$

where $P_0 = 1 - \lambda\alpha$ and $U(s, z)$ is defined by (7). Thus

$$\begin{aligned} \Omega(s) = (1 - \lambda\alpha) & \left\{ 1 + \frac{\lambda}{s} \int_{\gamma(s)}^1 \right. \\ & \cdot \exp \left[- \int_u^1 \frac{dv}{v - \Psi(s + \lambda(1 - v))} \right] \\ & \cdot \left. \left[\frac{u - 1}{u - \Psi(\lambda(1 - u))} - \frac{u - 1}{u - \Psi(s + \lambda(1 - u))} \right] du \right\}, \end{aligned} \quad (63)$$

whence (53) follows by integrating by parts and by using the fact that $\gamma(s)$ satisfies (54) in z .

If α_{n+1} is finite, then W_n is also finite and can be expressed by

$$\int_0^{\infty} x^r dW_j(x) = (-1)^r \left(\frac{d^r \Omega_j(s)}{ds^r} \right)_{s=0} \quad (64)$$

for $r = 1, 2, \dots, n$ and $j = 1, 2, \dots$, and by

$$U_{jk} = (-1)^j \left(\frac{\partial^{j+k} U(s, z)}{\partial s^j \partial z^k} \right)_{s=0, z=1} \quad (65)$$

for $j + k \leq n$. By using the following formulas we obtain (34) and (36):

$$\int_0^{\infty} x dW_j(x) = \frac{\alpha(j - 1)}{2 - \lambda\alpha}, \quad (66)$$

$$\int_0^{\infty} x^2 dW_j(x) = \frac{2\alpha^2(j-1)(j-2)}{(2-\lambda\alpha)(3-2\lambda\alpha)} + \frac{(6-\lambda\alpha)\alpha_2(j-1)}{(2-\lambda\alpha)^2(3-2\lambda\alpha)}, \quad (67)$$

and further $U_{00} = \lambda\alpha$, $U_{10} = \lambda\alpha_2/2$, $U_{20} = \lambda\alpha_3/3$,

$$U_{01} = \lambda\alpha + \frac{\lambda^2\alpha_2}{2(1-\lambda\alpha)},$$

$$U_{11} = \frac{\lambda^2\alpha_3}{6} + \frac{\lambda\alpha_2}{2} + \frac{\lambda^3\alpha_2^2}{4(1-\lambda\alpha)},$$

and

$$U_{02} = \frac{\lambda^3\alpha_3 + 3\lambda^2\alpha_2}{3(1-\lambda\alpha)} + \frac{\lambda^4\alpha_2^2}{2(1-\lambda\alpha)^2}.$$

(iii) *Service in Reverse Order of Arrival.* The case of exponentially distributed service times was investigated by E. Vulot¹¹ and the general case by J. Riordan¹² and D. M. G. Wishart.¹³ Now we shall prove

Theorem 4: If $\lambda\alpha < 1$, if the process is stationary and if service is in reverse order of arrival, then the distribution function of the delay of a call for $x \geq 0$ is given by

$$W(x) = (1 - \lambda\alpha) + \lambda \sum_{j=1}^{\infty} e^{-\lambda x} \frac{(\lambda x)^{j-1}}{j!} \int_0^x [1 - H_j(u)] du \quad (68)$$

where $H_j(x)$ denotes the j th iterated convolution of $H(x)$ with itself.

Proof: Denote by $G(x)$ the probability that the length of a busy period in the queueing process considered is $\leq x$. $G(x)$ is given by (23). Then we can write that

$$W(x) = P_0 + (1 - P_0) \sum_{k=0}^{\infty} \left[\int_0^x e^{-\lambda u} \frac{(\lambda u)^k}{k!} dH^*(u) \right] * G_k(x) \quad (69)$$

where $G_k(x)$ denotes the k th iterated convolution of $G(x)$ with itself; $G_0(x) = 1$ if $x \geq 0$ and $G_0(x) = 0$ if $x < 0$; $H^*(x)$ is defined by (18) and $P_0 = 1 - \lambda\alpha$. For, if an arriving call finds the line free, which has probability P_0 , then its service starts without delay; if the line is busy, which has probability $1 - P_0$, then its delay is composed of the remaining holding time of the call being served, which has the distribution function $H^*(x)$, and if during this time interval k new calls join the queue, then there is an additional delay that has the same distribution function as the total length of k independent busy periods. Thus we obtain (69). Forming the Laplace-Stieltjes transform of (69) we obtain

$$\begin{aligned} \Omega(s) &= P_0 + \frac{(1 - P_0)}{\alpha} \sum_{k=0}^{\infty} [\gamma(s)]^k \int_0^{\infty} e^{-(\lambda+s)x} \frac{(\lambda x)^k}{k!} [1 - H(x)] dx \\ &= (1 - \lambda\alpha) + \lambda \frac{1 - \Psi(s + \lambda(1 - \gamma(s)))}{s + \lambda[1 - \gamma(s)]} \end{aligned} \tag{70}$$

where $\gamma(s)$ is defined by (22). Since $\gamma(s)$ satisfies (21) in z , we get from (70) that

$$\Omega(s) = (1 - \lambda\alpha) + \frac{\lambda[1 - \gamma(s)]}{s + \lambda[1 - \gamma(s)]}. \tag{71}$$

By using Lagrange's expansion (cf. Ref. 3, p. 132) we obtain

$$\begin{aligned} \Omega(s) &= (1 - \lambda\alpha) + \frac{\lambda}{\lambda + s} \\ &\quad + s \sum_{j=1}^{\infty} \frac{(-1)^j \lambda^j}{j!} \frac{d^{j-1}}{ds^{j-1}} \left(\frac{[\Psi(s + \lambda)]^j}{(s + \lambda)^2} \right), \end{aligned} \tag{72}$$

whence (68) follows by inversion.

If α_{n+1} is finite then W_n is also finite and we have for $n = 2, 3, \dots$ that

$$W_n = \sum_{\nu=1}^n \frac{(n - 2 + \nu)! \lambda^\nu}{(n - 1)! (1 - \lambda\alpha)^{n-1+\nu}} Y_{n,\nu} \tag{73}$$

where $Y_{n\nu}$ is defined by (26). If $n = 1$ then W_n is given by (34). For,

$$W_n = (-1)^n \left(\frac{d^n \Omega(s)}{ds^n} \right)_{s=0} \quad (n = 0, 1, 2, \dots). \tag{74}$$

If we use the notation $u = s + \lambda[1 - \gamma(s)]$ and $s = u - \lambda[1 - \Psi(u)]$, then we can write that

$$\Omega(s) = (1 - \lambda\alpha) + \lambda \frac{1 - \Psi(u)}{u}, \tag{75}$$

whence by using Bürmann's theorem (cf. Appendix) we obtain for $n = 2, 3, \dots$ that

$$W_n = \frac{(-1)^{n-1}}{n - 1} \left(\frac{d^n}{du^n} \left[\frac{1}{1 - \lambda \frac{1 - \Psi(u)}{u}} \right]^{n-1} \right)_{u=0} \tag{76}$$

and the n th derivative can be calculated by Faa di Bruno's formula (cf. Appendix).

In this case W_1 is given by (34), W_2 by (37) and

$$W_3 = \frac{\lambda\alpha_4}{4(1-\lambda\alpha)^3} + \frac{3\lambda^2\alpha_2\alpha_3}{2(1-\lambda\alpha)^4} + \frac{3\lambda^3\alpha_2^3}{2(1-\lambda\alpha)^5}. \quad (77)$$

Remark. If $P_0(t)$ denotes the probability that the line is free at time t given that it was free at time $t = 0$, then we can write that

$$W(x) = 1 - [P_0(x) - P_0(\infty)] \quad (78)$$

where $P_0(\infty) = 1 - \lambda\alpha$.

If $G^*(x)$ denotes the probability that the length of a busy period is $\leq x$ for the dual process $[H(x), F(x), 1]$, i.e., when the interarrival times and holding times are interchanged, then we can write that

$$W(x) = 1 - [G^*(\infty) - G^*(x)] \quad (79)$$

where $G^*(\infty) = \lambda\alpha$.

(iv) *At Extreme Case.* Suppose that in the stationary process the service of a particularly chosen call starts when and only when no other calls are in the system, i.e., its service is delayed until it becomes the only call in the system. Denote by $W^*(x)$ the distribution function of the delay of this particular call.

Theorem 5: If $\lambda\alpha < 1$, if the process is stationary, and if a particularly chosen call will be served only when no other calls are in the system, then the distribution function of the delay of this call is given by $W^*(0) = 1 - \lambda\alpha$ and for $x > 0$

$$\frac{dW^*(x)}{dx} = (1 - \lambda\alpha)\lambda[1 - G(x)], \quad (80)$$

where $G(x)$ is the distribution function of the length of a busy period and is given by (23).

Proof: Denote by $G_n(x)$ the n th iterated convolution of $G(x)$ with itself. Now we have

$$W^*(x) = P_0 + \sum_{j=1}^{\infty} \sum_{k=1}^{\infty} \left[\int_0^x e^{-\lambda u} \frac{(\lambda u)^k}{k!} dP_j(u) \right] * G_{j+k}(x). \quad (81)$$

For, if the particularly chosen call finds the line to be free, then its service starts without delay; if it finds j calls ($j = 1, 2, \dots$) in the system and if during the remaining part of the current service k ($k = 0, 1, \dots$) more calls arrive, then its delay is composed of the remaining holding time of the call being served at its arrival and an additional delay which has the same distribution as the sum of $j + k$ mutually independent random variables each of which has the same distribution as a busy period.

Denote by $\Omega^*(s)$ the Laplace-Stieltjes transform of $W^*(x)$. By (81)

$$\begin{aligned} \Omega^*(s) &= P_0 + \sum_{j=1}^{\infty} \Pi_j (s + \lambda - \lambda\gamma(s))[\gamma(s)]^j \\ &= P_0 + U(s + \lambda - \lambda\gamma(s), \gamma(s)) = P_0 \left[1 - \lambda \frac{1 - \gamma(s)}{s} \right], \end{aligned} \tag{82}$$

where $P_0 = 1 - \lambda\alpha$ and $\gamma(s)$ is given by (22). We obtain (80) by inversion.

If α_{n+1} is finite then

$$W_n^* = \int_0^{\infty} x^n dW^*(x) \tag{83}$$

is also finite and is given by

$$W_n^* = \frac{\lambda(1 - \lambda\alpha)\Gamma_{n+1}}{n + 1} \quad (n = 1, 2, \dots) \tag{84}$$

where Γ_{n+1} is defined by (25). This follows immediately from (82). In particular we have

$$W_1^* = \frac{\lambda\alpha_2}{2(1 - \lambda\alpha)^2}, \tag{85}$$

$$W_2^* = \frac{\lambda\alpha_3}{3(1 - \lambda\alpha)^3} + \frac{\lambda^2\alpha_2^2}{(1 - \lambda\alpha)^4}, \tag{86}$$

$$W_3^* = \frac{\lambda\alpha_4}{4(1 - \lambda\alpha)^4} + \frac{5\lambda^2\alpha_2\alpha_3}{2(1 - \lambda\alpha)^5} + \frac{15\lambda^3\alpha_2^3}{4(1 - \lambda\alpha)^6}. \tag{87}$$

APPENDIX

A.1 *Bürmann's Theorem*

(Cf. Ref. 3, p. 128.) Suppose that the first N derivatives of $f(z)$ and the first $N - 1$ derivatives of $g(z)$ exist at $z = 0$. If $s = u/g(u)$ and $g(0) \neq 0$, then

$$f(u) = f(0) + \sum_{n=1}^N \frac{s^n}{n!} \left(\frac{d^{n-1}f'(v)[g(v)]^n}{dv^{n-1}} \right)_{v=0} + o(s^N). \tag{88}$$

A.2 *Faa di Bruno's Formula*

(Cf. Ref. 14, p. 33.) If $z = f(y)$ where $y = g(x)$, then the n th derivative of $z = f(g(x))$ with respect to x at $x = 0$ is given by the following

TABLE I

n	ν	j_1	j_2	j_3	j_4	j_5	C_{j_1, j_2, \dots, j_n}
1	1	1					$1/2$
2	1	0	1				$1/3$
2	2	2	0				$1/4$
3	1	0	0	1			$1/4$
3	2	1	1	0			$1/2$
3	3	3	0	0			$1/6$
4	1	0	0	0	1		$1/5$
4	2	1	0	1	0		$1/2$
4	2	0	2	0	0		$1/3$
4	3	2	1	0	0		$1/2$
4	4	4	0	0	0		$1/6$
5	1	0	0	0	0	1	$1/6$
5	2	1	0	0	1	0	$1/2$
5	2	0	1	1	0	0	$5/6$
5	3	2	0	1	0	0	$5/6$
5	3	1	2	0	0	0	$5/6$
5	4	3	1	0	0	0	$5/6$
5	5	5	0	0	0	0	$1/32$

formula

$$\left(\frac{d^n f(g(x))}{dx^n} \right)_{x=0} = \sum_{\nu=1}^n Y_{n,\nu} \left(\frac{d^\nu f(y)}{dy^\nu} \right)_{y=g(0)} \quad (89)$$

where

$$Y_{n,\nu} = \sum_{\substack{j_1+j_2+\dots+j_n=\nu \\ j_1+2j_2+\dots+nj_n=n}} \frac{n! [g^{(1)}(0)]^{j_1} [g^{(2)}(0)]^{j_2} \dots [g^{(n)}(0)]^{j_n}}{j_1! j_2! \dots j_n! (1!)^{j_1} (2!)^{j_2} \dots (n!)^{j_n}} \quad (90)$$

provided that the derivatives in question exist.

For $n \leq 5$ Table I contains all the n -tuples (j_1, j_2, \dots, j_n) satisfying the requirements $j_1 + j_2 + \dots + j_n = \nu$ and $j_1 + 2j_2 + \dots + nj_n = n$, and in addition the coefficients

$$C_{j_1, j_2, \dots, j_n} = \frac{n!}{j_1! j_2! \dots j_n! (2!)^{j_1} (3!)^{j_2} \dots ((n+1)!)^{j_n}}, \quad (91)$$

which we need in using formula (26).

REFERENCES

1. Wishart, D. M. G., An Application of Ergodic Theorems in the Theory of Queues, Proceedings of the Fourth Berkeley Symposium on Mathematical Statistics and Probability, Univ. of California Press, Berkeley and Los Angeles, Vol. 2, 1961, pp. 581-592.
2. Takács, L., *Introduction to the Theory of Queues*, Oxford University Press, New York, 1962.
3. Whittaker, E. T., and Watson, G. N., *A Course of Modern Analysis*, Cambridge University Press, Cambridge, 1952.

4. Pollaczek, F., Über eine Aufgabe der Wahrscheinlichkeitstheorie I-II. *Mathematische Zeitschrift* **32**, 1930, pp. 64-100 and 729-750.
5. Khintchine, A. Y., Mathematical Theory of Stationary Queues (in Russian), *Matem. Sbornik*, **39**, No. 4, 1932, pp. 73-84.
6. Lindley, D. V., The Theory of Queues with a Single Server. *Proc. Cambridge Phil. Soc.* **48**, 1952, pp. 277-289.
7. Beneš, V. E., On Queues with Poisson Arrivals, *Ann. Math. Statist.* **28**, 1957, pp. 670-77.
8. Takács, L., Delay Distributions for Simple Trunk Groups with Recurrent Input and Exponential Service Times, *B.S.T.J.*, **41**, January, 1962, pp. 311-320.
9. Burke, P. J., Equilibrium Delay Distribution for One Channel with Constant Holding Time, Poisson Input and Random Service, *B.S.T.J.*, **38**, July, 1959, pp. 1021-1031.
10. Kingman, J. F. C., On Queues in Which Customers Are Served in Random Order, *Proc. Cambridge Phil. Soc.* **58**, 1962, pp. 79-91.
11. Vulot, E., Delais d'attente des appels téléphoniques dans l'ordre inverse de leur arrivée, *Comptes Rendus Acad. Sci. Paris* **238**, 1954, pp. 1188-1189.
12. Riordan, J., Delays for Last-Come First-Served Service and the Busy Period. *B.S.T.J.*, **40**, May, 1961, pp. 785-793.
13. Wishart, D. M. G., Queuing Systems in Which the Discipline Is 'Last-Come, First-Served', *Operations Research* **8**, 1960, pp. 591-599.
14. Jordan, Ch., *Calculus of Finite Differences*, Budapest, 1939. Second ed., Chelsea, New York, 1947.

A Single-Server Queue with Feedback

By L. TAKÁCS

(Manuscript received December 14, 1962)

Let us suppose that customers arrive at a counter in accordance with a Poisson process of density λ . The customers are served by a single server in order of arrival. The service times are identically distributed, mutually independent, positive random variables with distribution function $H(x)$. Suppose that after being served each customer either immediately joins the queue again with probability p or departs permanently with probability q ($p + q = 1$). In this paper we shall determine for a stationary process the distribution of the queue size as well as the Laplace-Stieltjes transform and the first two moments of the distribution function of the total time spent in the system by a customer.

I. INTRODUCTION

Although the problems discussed in this paper arose in the theory of telephone traffic, we use the terminology of queues. Thus instead of calls and holding times we shall speak about customers and service times respectively.

Let us suppose that in the time interval $(0, \infty)$ customers arrive at a counter in accordance with a Poisson process of density λ . Denote by τ_n ($n = 1, 2, \dots$) the arrival time of the n th customer. Then the inter-arrival times $\tau_{n+1} - \tau_n$ ($n = 0, 1, \dots$; $\tau_0 = 0$) are identically distributed, mutually independent random variables with distribution function

$$F(x) = \begin{cases} 1 - e^{-\lambda x} & \text{if } x \geq 0, \\ 0 & \text{if } x < 0. \end{cases} \quad (1)$$

The customers are served by a single server in order of arrival. The server is idle if and only if there is no customer in the system. The service times are supposed to be identically distributed, mutually independent, positive random variables with distribution function $H(x)$, and independent of the input process. Suppose that after being served each customer either immediately joins the queue again with probability p or goes away permanently with probability q where $p + q = 1$. The event

that a customer returns is independent of any other event involved and, in particular, independent of the number of his previous returns.

The process defined above is said to be of type $[F(x), H(x), p]$. If $p = 0$, then there is no feedback.

Let us denote by $\xi(t)$ the queue size at time t , that is, $\xi(t)$ is the number of customers (either waiting or being served) in the system at time t . Let ξ_n denote the queue size immediately before the arrival of the n th customer, that is, $\xi_n = \xi(\tau_n - 0)$. Denote by θ_n the total time spent in the system by the n th customer.

We are interested in finding the distribution of $\xi(t)$ for a stationary process, i.e., when $\xi(t)$ has the same distribution for all $t \geq 0$ and the distribution of θ_n for a stationary process, i.e., when θ_n has the same distribution for every $n = 1, 2, \dots$.

It is easy to prove that the limiting distribution $\lim_{t \rightarrow \infty} \mathbf{P}\{\xi(t) = k\}$ ($k = 0, 1, \dots$) exists and is independent of the initial state if and only if $\xi(t)$ has a stationary distribution and the limiting distribution is identical with the stationary distribution. Similarly the limiting distribution $\lim_{n \rightarrow \infty} \mathbf{P}\{\theta_n \leq x\}$ exists and is independent of the initial state if and only if θ_n has a stationary distribution and the limiting distribution is identical with the stationary distribution.

Throughout this paper we use the following notation:

$$\Psi(s) = \int_0^{\infty} e^{-sx} dH(x) \quad (\Re(s) \geq 0) \quad (2)$$

for the Laplace-Stieltjes transform of $H(x)$,

$$\alpha_r = \int_0^{\infty} x^r dH(x) \quad (r = 0, 1, \dots) \quad (3)$$

for the r th moment of $H(x)$, and

$$\alpha = \int_0^{\infty} x dH(x) \quad (4)$$

for the average service time, i.e., $\alpha = \alpha_1$.

Further denote by $H^*(x)$ the distribution function of the total service time of a customer. We have

$$H^*(x) = g \sum_{k=1}^{\infty} p^{k-1} H_k(x) \quad (5)$$

where $H_k(x)$ denotes the k th iterated convolution of $H(x)$ with itself. For gp^{k-1} ($k = 1, 2, \dots$) is the probability that a customer joins the

queue k times, and if he joins k times, his total service time is equal to the sum of k mutually independent random variables each of which has the distribution function $H(x)$. If we introduce the notation

$$\Psi^*(s) = \int_0^\infty e^{-sx} dH^*(x) \quad (\Re(s) \geq 0), \quad (6)$$

then by (5) we obtain that

$$\Psi^*(s) = q \sum_{k=1}^{\infty} p^{k-1} [\Psi(s)]^k = \frac{q\Psi(s)}{1 - p\Psi(s)}. \quad (7)$$

Let

$$\alpha_r^* = \int_0^\infty x^r dH^*(x) \quad (r = 0, 1, \dots) \quad (8)$$

and $\alpha^* = \alpha_1^*$. By (7), $\alpha^* = \alpha/q$,

$$\alpha_2^* = \frac{\alpha_2}{q} + \frac{2p\alpha_1^2}{q^2}, \quad (9)$$

and in general α_r^* can be obtained by the following recurrence formula

$$\alpha_r^* = \alpha_r + \frac{p}{q} \sum_{j=1}^r \binom{r}{j} \alpha_j \alpha_{r-j}^*. \quad (10)$$

II. THE STATIONARY DISTRIBUTION OF THE QUEUE SIZE

If we know the stationary distribution of the queue size for a process of type $[F(x), H(x), 0]$, then that for a process of type $[F(x), H(x), p]$ can be obtained immediately.

Theorem 1: If $\lambda\alpha < q$, then the process $\{\xi(t), 0 \leq t < \infty\}$ has a unique stationary distribution $\mathbf{P}\{\xi(t) = j\} = P_j^*$ ($j = 0, 1, \dots$) and for $|z| \leq 1$

$$U^*(z) = \sum_{j=0}^{\infty} P_j^* z^j = \left(1 - \frac{\lambda\alpha}{q}\right) \frac{q(1-z)\Psi(\lambda(1-z))}{(q + pz)\Psi(\lambda(1-z)) - z}. \quad (11)$$

If $\lambda\alpha \geq q$, then a stationary distribution does not exist.

Proof: To find the distribution of the queue size we may assume without loss of generality that the customers join the queue only once and are served in one stretch; however, their service time is equal to the total service time that they would have if they were served in the original manner. Accordingly the distribution of the queue size for the process of type $[F(x), H(x), p]$ is the same as for the process of type

$$[F(x), H^*(x), 0].$$

For the latter a stationary distribution $\{P_j^*\}$ exists if and only if $\lambda\alpha^* < 1$, that is, $\lambda\alpha < q$, and is given by A. Y. Khintchine's formula:

$$U^*(z) = \sum_{j=0}^{\infty} P_j^* z^j = \frac{(1 - \lambda\alpha^*)(1 - z)\Psi^*(\lambda(1 - z))}{\Psi^*(\lambda(1 - z)) - z}. \quad (12)$$

(Cf., e.g., Ref. 1 or 2.) This proves (11).

Remark 1. Denote by B_r^* ($r = 0, 1, \dots$) the r th binomial moment of $\{P_j^*\}$, that is,

$$B_r^* = \sum_{j=r}^{\infty} \binom{j}{r} P_j^*. \quad (13)$$

If α_{r+1} is finite, then B_r^* is also finite. We have $B_0^* = 1$,

$$B_1^* = \frac{\lambda^2 \alpha_2^*}{2(1 - \lambda\alpha_1^*)} + \lambda\alpha_1^* = \frac{\lambda^2 \alpha_2 + 2\lambda\alpha_1(1 - \lambda\alpha_1)}{2(q - \lambda\alpha_1)} \quad (14)$$

and for $r = 2, 3, \dots$

$$B_r^* = \frac{\lambda^r}{r!} \left[\sum_{\nu=r}^r \frac{\lambda^{\nu} \nu!}{(1 - \lambda\alpha^*)^{\nu}} Y_{r,\nu} + \frac{r}{\lambda} \sum_{\nu=r-1}^{r-1} \frac{\lambda^{\nu} \nu!}{(1 - \lambda\alpha^*)^{\nu}} Y_{r-1,\nu} \right] \quad (15)$$

where

$$Y_{n,\nu} = \sum_{\substack{j_1+j_2+\dots+j_n=\nu \\ j_1+2j_2+\dots+nj_n=n}} \frac{n! \alpha_2^{*j_1} \alpha_3^{*j_2} \dots \alpha_{n+1}^{*j_n}}{j_1! j_2! \dots j_n! (2!)^{j_1} (3!)^{j_2} \dots ((n+1)!)^{j_n}} \quad (16)$$

The proof of (14) and (15) can be found in Ref. 3.

III. THE STATIONARY PROCESS

Let us denote by χ_n the time needed to complete the current service (if any) at the instant $t = \tau_n - 0$, i.e., immediately before the arrival of the n th customer. If $\xi_n = 0$, then $\chi_n = 0$. It is easy to see that the vector sequence $\{\xi_n, \chi_n; n = 1, 2, \dots\}$ is a Markov sequence. We shall prove that if $\lambda\alpha < q$, then $\{\xi_n, \chi_n\}$ has a unique stationary distribution, whereas if $\lambda\alpha \geq q$, then a stationary distribution does not exist. For a stationary sequence $\{\xi_n, \chi_n\}$ introduce the following notation:

$$P_j = \mathbf{P}\{\xi_n = j\} \quad (j = 0, 1, \dots), \quad (17)$$

$$P_j(x) = \mathbf{P}\{\chi_n \leq x, \xi_n = j\} \quad (j = 1, 2, \dots) \quad (18)$$

and

$$\Pi_j(s) = \int_0^{\infty} e^{-sx} dP_j(x) \quad (j = 1, 2, \dots). \quad (19)$$

Theorem 2: If $\lambda\alpha < q$, then the Markov sequence $\{\xi_n, \chi_n; n = 1, 2, \dots\}$ has a unique stationary distribution, for which $P_j = P_j^*$ defined by (11) and

$$\begin{aligned}
 U(s,z) &= \sum_{j=1}^{\infty} \Pi_j(s)z^j \\
 &= \left(1 - \frac{\lambda\alpha}{q}\right) \frac{\lambda z(1-z)[\Psi(s) - \Psi(\lambda(1-z))]}{[z - (q + pz)\Psi(\lambda(1-z))][s - \lambda(1-z)]}.
 \end{aligned}
 \tag{20}$$

If $\lambda\alpha \geq q$, then $\{\xi_n, \chi_n\}$ has no stationary distribution.

Proof. First consider the process of type $[F(x), H^*(x), 0]$. It is proved in Ref. 3 that in this case $\{\xi_n, \chi_n\}$ has a unique stationary distribution if and only if $\lambda\alpha < q$. Namely $\mathbf{P}\{\xi_n = j\} = P_j^*$ ($j = 0, 1, \dots$) given by (11) and the generating function corresponding to (20) is given by

$$U^*(s,z) = \frac{(1 - \lambda\alpha)\lambda z(1-z)[\Psi^*(s) - \Psi^*(\lambda(1-z))]}{[z - \Psi^*(\lambda(1-z))][s - \lambda(1-z)]}. \tag{21}$$

The distribution of the queue size is the same for both the types

$$[F(x), H(x), p] \quad \text{and} \quad [F(x), H^*(x), 0].$$

The only difference between the process of type $[F(x), H(x), p]$ and $[F(x), H^*(x), 0]$ is that in the latter χ_n the remaining length of the current service at the arrival of the n th customer is replaced by the remaining part of the total service time of the customer just being served at the arrival of the n th customer. The time added to χ_n is independent of the queue size and has the distribution function

$$\hat{H}(x) = q \sum_{k=0}^{\infty} p^k H_k(x), \tag{22}$$

because the probability that a departing customer will join the queue k more times is qp^k and in this case the additional total service time has the distribution function $H_k(x)$; $H_0(x) = 1$ if $x \geq 0$ and $H_0(x) = 0$ if $x < 0$. The Laplace-Stieltjes transform of (22) is

$$\hat{\Psi}(s) = \int_0^{\infty} e^{-sx} d\hat{H}(x) = q \sum_{k=0}^{\infty} p^k [\Psi(s)]^k = \frac{q}{1 - p\Psi(s)}. \tag{23}$$

Accordingly we have

$$U^*(s,z) = U(s,z)\hat{\Psi}(s), \tag{24}$$

whence (20) follows.

Remark 2. If $\chi(t)$ denotes the time needed to complete the current

service (if any) at time t ($\chi(t) = 0$ if $\xi(t) = 0$), then the vector process $\{\xi(t), \chi(t); 0 \leq t < \infty\}$ is a Markov process. $\{\xi(t), \chi(t)\}$ has a stationary distribution if and only if $\lambda\alpha < q$ and it agrees with the stationary distribution of $\{\xi_n, \chi_n\}$.

IV. THE STATIONARY DISTRIBUTION OF θ_n

If the joint distribution of ξ_n and χ_n is known, then the distribution of θ_n is determined uniquely. If $\{\xi_n, \chi_n\}$ has a stationary distribution, then every θ_n ($n = 1, 2, \dots$) has the same distribution. In case of a stationary process let

$$\mathbf{E}\{e^{-s\theta_n}\} = \Phi(s) \quad (\Re(s) \geq 0). \quad (25)$$

Theorem 3: If $\lambda\alpha < q$, then θ_n has a unique stationary distribution $\mathbf{P}\{\theta_n \leq x\}$, which is given by the following Laplace-Stieltjes transform

$$\Phi(s) = q \sum_{k=1}^{\infty} p^{k-1} U_k(s, 1) \quad (\Re(s) \geq 0), \quad (26)$$

where

$$U_1(s, z) = P_0 \Psi(s + \lambda(1 - z)) + U(s + \lambda(1 - z), (q + pz) \Psi(s + \lambda(1 - z))) \quad (27)$$

for $\Re(s) \geq 0$ and $|z| \leq 1$, $P_0 = 1 - \lambda\alpha/q$, $U(s, z)$ is defined by (20), and

$$U_{k+1}(s, z) = \Psi(s + \lambda(1 - z)) U_k(s, (q + pz) \Psi(s + \lambda(1 - z))) \quad (28)$$

for $k = 1, 2, \dots$.

Proof: The probability that a customer joins the queue exactly k times (including the original arrival) is qp^{k-1} ($k = 1, 2, \dots$). Denote by $\theta_n^{(k)}$ the total time spent in the system by the n th customer until his k th departure (if he joins the queue at least k times). Denote by $\zeta_n^{(k)}$ the queue size immediately after the k th departure of the n th customer. Let

$$U_k(s, z) = \mathbf{E}\{\exp[-s\theta_n^{(k)}] z^{\zeta_n^{(k)}}\} \quad (k = 1, 2, \dots). \quad (29)$$

We can easily see that for a stationary sequence $\{\xi_n, \chi_n\}$

$$\begin{aligned} U_1(s, z) &= P_0 \Psi(s + \lambda(1 - z)) \\ &\quad + \sum_{j=1}^{\infty} \Pi_j(s + \lambda(1 - z)) (q + pz)^j [\Psi(s + \lambda(1 - z))]^j \\ &= P_0 \Psi(s + \lambda(1 - z)) \\ &\quad + U(s + \lambda(1 - z), (q + pz) \Psi(s + \lambda(1 - z))) \end{aligned} \quad (30)$$

where $P_0 = 1 - \lambda\alpha/q$ and $U(s, z)$ is defined by (20). Now we shall prove that (28) holds for $k = 1, 2, \dots$. Under the conditions $\zeta_n^{(k)} = j$, $\theta_n^{(k)} = x$ and that after the k th service the n th customer joins the queue again, the difference $\theta_n^{(k+1)} - \theta_n^{(k)}$ is equal to the length of $j + 1$ services, the distribution function of which is $H_{j+1}(x)$. If $\theta_n^{(k+1)} - \theta_n^{(k)} = y$, then $\zeta_n^{(k+1)}$ is equal to the sum of two independent random variables: the first is the number of customers arriving at the counter during the time interval of length y , which has a Poisson distribution with parameter λy , and the second is the number of returning customers, which has a Bernoulli distribution with parameters j and p . Thus

$$\begin{aligned} \mathbf{E}\{\exp[-s\theta_n^{(k+1)}]z^{\zeta_n^{(k+1)}} \mid \zeta_n^{(k)} = j, \theta_n^{(k)} = x\} &= e^{-sx}(q + pz)^j. \quad (31) \\ \int_0^\infty \exp[-sy - \lambda(1 - z)y] dH_{j+1}(y) & \\ &= e^{-sx}(q + pz)^j[\Psi(s + \lambda(1 - z))]^{j+1}, \end{aligned}$$

or

$$\begin{aligned} \mathbf{E}\{\exp[-s\theta_n^{(k+1)}]z^{\zeta_n^{(k+1)}} \mid \zeta_n^{(k)}, \theta_n^{(k)}\} & \\ = \Psi(s + \lambda(1 - z)) \exp[-s\theta_n^{(k)}] \cdot [(q + pz)\Psi(s + \lambda(1 - z))]^{\zeta_n^{(k)}} & \quad (32) \end{aligned}$$

and unconditionally

$$U_{k+1}(s, z) = \Psi(s + \lambda(1 - z))U_k(s, (q + pz)\Psi(s + \lambda(1 - z))) \quad (33)$$

which proves (28). It is to be noted that stationarity has been used only in the determination of $U_1(s, z)$. The recurrence relation (28) is valid for any process. Finally,

$$\mathbf{E}\{\exp[-s\theta_n]\} = q \sum_{k=1}^{\infty} p^{k-1} \mathbf{E}\{\exp[-s\theta_n^{(k)}]\} = q \sum_{k=1}^{\infty} p^{k-1} U_k(s, 1) \quad (34)$$

which was to be proved.

V. THE MOMENTS OF θ_n

Although it seems very complicated to find a closed formula for $\Phi(s)$, the moments of θ_n can be determined explicitly. We shall prove that $\Phi(s) = \Phi(s, 1)$ where $\Phi(s, z)$ satisfies a functional equation. This observation makes it possible to find explicit formulas for the moments of θ_n .

Theorem 4: If $\lambda\alpha < q$ and θ_n has a stationary distribution, then

$$\mathbf{E}\{\theta_n\} = \frac{\lambda\alpha_2 + 2\alpha_1(1 - \lambda\alpha_1)}{2(q - \lambda\alpha_1)} \quad (35)$$

provided that α_2 is finite, and

$$\mathbf{E}\{\theta_n^2\} = \frac{q^2 - 2q}{6(q - \lambda\alpha_1)^2[q^2 - q(2 + \lambda\alpha_1) + \lambda\alpha_1]} \cdot \{2q[6\lambda\alpha_1^3 - 6\alpha_1^2 - 6\lambda\alpha_1\alpha_2 + 3\alpha_2 + \lambda\alpha_3] - [12\lambda\alpha_1^3 - 12\alpha_1^2 - 6\lambda\alpha_1\alpha_2 + 2\lambda^2\alpha_1\alpha_3 - 3\lambda^2\alpha_2^2]\}, \quad (36)$$

provided that α_3 is finite.

Proof: Let

$$\Phi(s, z) = q \sum_{k=1}^{\infty} p^{k-1} U_k(s, z) \quad (37)$$

where $U_k(s, z)$ is defined by (27) and (28). By using the recurrence relation (28) we obtain that

$$\Phi(s, z) = qU_1(s, z) + p\Psi(s + \lambda(1 - z))\Phi(s, (q + pz)\Psi(s + \lambda(1 - z))). \quad (38)$$

Let

$$\Phi_{ij} = \left(\frac{\partial^{i+j}\Phi(s, z)}{\partial s^i \partial z^j} \right)_{s=0, z=1}. \quad (39)$$

If we form Φ_{ij} by (38) for $i + j = r$ ($i = 0, 1, \dots, r$), then we obtain $r + 1$ linear equations for the determination of Φ_{ij} . These equations can be solved successively for $r = 1, 2, \dots$. By (26)

$$\mathbf{E}\{\theta_n^r\} = (-1)^r \Phi_{r0} \quad (r = 0, 1, \dots) \quad (40)$$

for a stationary process.

VI. A PARTICULAR CASE

If we suppose, in particular, that the service times have an exponential distribution

$$H(x) = \begin{cases} 1 - e^{-\mu x} & \text{for } x \geq 0, \\ 0 & \text{for } x < 0, \end{cases} \quad (41)$$

then $H^*(x) = 1 - e^{-\mu q x}$ for $x \geq 0$, $\Psi(s) = \mu/(\mu + s)$ and $\Psi^*(s) = \mu q/(\mu q + s)$. In this case (27) and (28) reduce to

$$U_{k+1}(s, z) = \frac{\mu}{\mu + s + \lambda(1 - z)} U_k \left(s, \frac{\mu(q + pz)}{\mu + s + \lambda(1 - z)} \right) \quad (42)$$

for $k = 0, 1, \dots$, where

$$U_0(s, z) = \left(1 - \frac{\lambda}{\mu q} \right) / \left(1 - \frac{\lambda z}{\mu q} \right). \quad (43)$$

By (42) we obtain that

$$U_k(s,1) = \frac{\left(1 - \frac{\lambda}{\mu q}\right)}{a_k(s) - b_k(s)} \quad (44)$$

where $a_k(s)$ and $b_k(s)$ ($k = 0, 1, \dots$) are given by the following matrix equation

$$\begin{pmatrix} a_k(s) \\ b_k(s) \end{pmatrix} = \begin{pmatrix} \frac{\mu + \lambda + s}{\mu}, & -q \\ \frac{\lambda}{\mu} & p \end{pmatrix}^k \begin{pmatrix} 1 \\ \frac{\lambda}{\mu q} \end{pmatrix}. \quad (45)$$

Now by (26) the Laplace-Stieltjes transform of $\mathbf{P}\{\theta_n \leq x\}$ is

$$\Phi(s) = \left(1 - \frac{\lambda}{\mu q}\right) \sum_{k=1}^{\infty} \frac{q p^{k-1}}{a_k(s) - b_k(s)}. \quad (46)$$

In this particular case

$$\mathbf{E}\{\theta_n\} = \frac{1}{\mu q - \lambda} \quad (47)$$

and

$$\mathbf{E}\{\theta_n^2\} = \frac{2\mu(2q - q^2)}{(\mu q - \lambda)^2[\mu(2q - q^2) - \lambda(1 - q)]}. \quad (48)$$

VII. ACKNOWLEDGMENTS

I am indebted to Dr. H. O. Pollak and Mr. J. P. Runyon, who called my attention to the problem discussed in this paper. I would like also to thank Dr. W. S. Brown for programming an IBM 7090 computer at the Bell Telephone Laboratories, Murray Hill, to determine the second moment of θ_n which is given by formula (36). It is most remarkable that Dr. Brown's method made it possible to use an electronic computer for performing algebraic operations.

APPENDIX—BY W. S. BROWN

Calculation of the Second Moment of θ_n

If $\lambda\alpha < q$ and if θ_n (the total time spent in the system by the n th customer) has a stationary distribution, then the moments of θ_n are determined by (38)–(40) as explained in Section V. The first moment can be calculated by hand without serious difficulty. The calculation of

the second moment, with the aid of an IBM 7090 computer,⁴ is described in this appendix.

It is convenient to replace z by $t = 1 - z$ and U_1 by $W = qU_1/(q - \lambda\alpha_1)$. Then the r th moment of θ_n is

$$\beta_r = (-1)^r \Phi_{r0} \quad (49)$$

where

$$\Phi_{ij} = \left[\left(\frac{\partial}{\partial s} \right)^i \left(\frac{\partial}{\partial t} \right)^j \Phi(s, t) \right]_{s=0, t=0}. \quad (50)$$

The function $\Phi(s, t)$ is implicitly defined by the equation

$$\Phi(s, t) = (q - \lambda\alpha_1)W(s, t) + p\psi(s + \lambda t)\Phi(s, \omega(s, t)) \quad (51)$$

where*

$$\psi(s) = \sum_{r=0}^{\infty} \frac{(-1)^r \alpha_r s^r}{r!} \quad (52)$$

$$\omega(s, t) = 1 - (1 - pt)\psi(s + \lambda t)$$

$$W(s, t) = \psi(s + \lambda t) + S(s + \lambda t, \lambda\omega(s, t))T(\omega(s, t))$$

with

$$S(x, y) = \frac{\psi(x) - \psi(y)}{x - y} \quad (53)$$

$$T(\omega) = \frac{\lambda\omega(1 - \omega)}{1 - \omega - (1 - p\omega)\psi(\lambda\omega)}.$$

This last pair of equations can be rewritten in the more useful form

$$S(x, y) = \sum_{r=0}^{\infty} \frac{(-1)^{r+1} \alpha_{r+1}}{(r+1)!} C_r(x, y) \quad (54)$$

$$T(\omega) = -\frac{\lambda(1 - \omega)}{q - \lambda(1 - p\omega)\varphi(\lambda\omega)}$$

where

$$C_r(x, y) = \frac{x^{r+1} - y^{r+1}}{x - y} = \sum_{k=0}^r x^k y^{r-k} \quad (55)$$

$$\varphi(x) = \frac{1 - \psi(x)}{x} = \sum_{r=0}^{\infty} \frac{(-1)^r \alpha_{r+1} x^r}{(r+1)!}.$$

* For convenience we have assumed that all of the service moments α_r [see (3)] are finite. However for the calculation of β_r it is clearly sufficient to require only the finiteness of α_{r+1} .

It is now clear that

$$\begin{aligned}\psi(0) &= \alpha_0 = 1 \\ S(0,0) &= -\alpha_1 \\ T(0) &= -\frac{\lambda}{q - \lambda\alpha_1} \\ W(0,0) &= \frac{q}{q - \lambda\alpha_1},\end{aligned}\tag{56}$$

so from (49)–(51)

$$\beta_0 = \Phi_{00} = \Phi(0,0) = 1\tag{57}$$

as is required by the definition of the *zerorth* moment.

Now suppose all of the quantities Φ_{ij} for $i + j < r$, where r is some positive integer, have been calculated and are expressed as rational functions of λ and p (or q) and the service moments α_k . Then by differentiation of (51) we can obtain a system of $r + 1$ linear equations in the $r + 1$ unknowns, Φ_{ij} with $i + j = r$. These equations will also contain the quantities Φ_{ij} with $i + j < r$, which can be replaced by their known values. The solutions of this linear system will again be rational functions of λ and p (or q) and the service moments α_k . Theoretically this procedure permits the calculation of arbitrarily many of the moments, but in practice the calculations are extremely lengthy. The reader may wish to try the first moment as an exercise.

We shall now outline the computation of the first two moments, β_1 and β_2 . The first step was to fix the time scale by setting λ equal to one. The next step was to express $\Phi(s,t)$ as a Taylor series with coefficients Φ_{ij} and expand (51) to second order in s and t . Since rational function operations were not yet available, all denominators had to be eliminated by suitable premultiplications. The result obtained by the computer was a gigantic polynomial in s , t , Φ_{00} , Φ_{01} , Φ_{10} , Φ_{02} , Φ_{11} , Φ_{20} , α_1 , α_2 , α_3 , and q . We chose to view it as a polynomial in s and t . Setting $\Phi_{00} = 1$, all the terms independent of both s and t vanished as expected.

The equations for Φ_{01} and Φ_{10} , obtained by setting the coefficients of t and s to zero, are

$$\begin{aligned}2A\Phi_{01} + B &= 0 \\ 2C\Phi_{10} + 2D\Phi_{01} + E &= 0\end{aligned}\tag{58}$$

where

$$\begin{aligned}
 A &= (q - \alpha_1)[q^2 - q(\alpha_1 + 2) + \alpha_1] \\
 B &= -(2\alpha_1^2 - 2\alpha_1 - \alpha_2)[q^2 - q(\alpha_1 + 2) + \alpha_1] \\
 C &= -q(q - \alpha_1) \\
 D &= -\alpha_1(q - 1)(q - \alpha_1) \\
 E &= (2\alpha_1^2 - 2\alpha_1 - \alpha_2)[(\alpha_1 + 1)q - \alpha_1].
 \end{aligned} \tag{59}$$

These expressions were factored by hand. From (58) we have

$$\begin{aligned}
 \Phi_{01} &= -\frac{B}{2A} \\
 \Phi_{10} &= \frac{BD - AE}{2AC},
 \end{aligned} \tag{60}$$

so, using (59)

$$\Phi_{01} = \Phi_{10} = \frac{2\alpha_1^2 - 2\alpha_1 - \alpha_2}{2(q - \alpha_1)}. \tag{61}$$

Thus the mean time spent in the system (waiting time plus service time) is

$$\beta_1 = -\Phi_{10} = \alpha_1 \left(\frac{1 - \alpha_1}{q - \alpha_1} \right) + \frac{\alpha_2}{2(q - \alpha_1)}. \tag{62}$$

In the absence of feedback ($p = 0$, $q = 1$) this reduces to

$$\beta_1^{(0)} = \alpha_1 + \frac{\alpha_2}{2(1 - \alpha_1)} \tag{63}$$

where the first term is the mean service time and the second term is the familiar expression⁵ for the mean waiting time in a single server system.

The equations for Φ_{02} , Φ_{11} , and Φ_{20} , obtained by setting the coefficients of t^2 , st , and s^2 to zero (and replacing Φ_{10} by Φ_{01}), are

$$\begin{aligned}
 A_{13}\Phi_{02} + B_1\Phi_{01} + C_1 &= 0 \\
 A_{22}\Phi_{11} + A_{23}\Phi_{02} + B_2\Phi_{01} + C_2 &= 0 \\
 A_{31}\Phi_{20} + A_{32}\Phi_{11} + A_{33}\Phi_{02} + B_3\Phi_{01} + C_3 &= 0
 \end{aligned} \tag{64}$$

where

$$\begin{aligned}
 A_{13} &= 6(q - \alpha_1)^2[q^3 - q^2(2\alpha_1 + 3) + q(\alpha_1^2 + 4\alpha_1 + 3) - \alpha_1(\alpha_1 + 2)] \\
 A_{22} &= -12(q - \alpha_1)^2[q^2 - q(\alpha_1 + 2) + \alpha_1] \\
 A_{23} &= -12\alpha_1(q - 1)(q - \alpha_1)^2(q - \alpha_1 - 1) \\
 A_{31} &= 6q(q - \alpha_1)^2 \\
 A_{32} &= 12\alpha_1(q - 1)(q - \alpha_1)^2 \\
 A_{33} &= 6\alpha_1^2(q - 1)(q - \alpha_1)^2
 \end{aligned} \tag{65}$$

and

$$\begin{aligned}
 B_1 &= 6(q - 1)(q - \alpha_1)^2[4q\alpha_1 - (2\alpha_1^2 + 4\alpha_1 + \alpha_2)] \\
 B_2 &= 12(q - 1)(q - \alpha_1)^2[2q\alpha_1 - (2\alpha_1^2 + 3\alpha_1 + \alpha_2)] \\
 B_3 &= -6(q - 1)(q - \alpha_1)^2(2\alpha_1^2 + 2\alpha_1 + \alpha_2)
 \end{aligned} \tag{66}$$

and finally

$$\begin{aligned}
 C_1 &= -2q^4(6\alpha_1^3 - 6\alpha_1^2 - 6\alpha_1\alpha_2 + 3\alpha_2 + \alpha_3) + 3q^3(8\alpha_1^4 - 8\alpha_1^2\alpha_2 \\
 &\quad - 8\alpha_1^2 - 6\alpha_1\alpha_2 + 2\alpha_1\alpha_3 - \alpha_2^2 + 6\alpha_2 + 2\alpha_3) - 2q^2(6\alpha_1^5 \\
 &\quad + 12\alpha_1^4 - 6\alpha_1^3\alpha_2 - 6\alpha_1^3 - 21\alpha_1^2\alpha_2 + 3\alpha_1^2\alpha_3 - 12\alpha_1^2 \\
 &\quad - 3\alpha_1\alpha_2^2 + 7\alpha_1\alpha_3 - 3\alpha_2^2 + 9\alpha_2 + 3\alpha_3) + q(12\alpha_1^5 \\
 &\quad - 18\alpha_1^3\alpha_2 + 2\alpha_1^3\alpha_3 - 3\alpha_1^2\alpha_2^2 - 12\alpha_1^2\alpha_2 + 10\alpha_1^2\alpha_3 \\
 &\quad - 12\alpha_1^2 - 9\alpha_1\alpha_2^2 + 12\alpha_1\alpha_2 + 10\alpha_1\alpha_3 - 6\alpha_2^2) \\
 &\quad + \alpha_1(6\alpha_1^2\alpha_2 - 2\alpha_1^2\alpha_3 + 3\alpha_1\alpha_2^2 - 6\alpha_1\alpha_2 - 4\alpha_1\alpha_3 + 3\alpha_2^2) \\
 C_2 &= 2q^3(12\alpha_1^4 - 18\alpha_1^3 - 12\alpha_1^2\alpha_2 + 6\alpha_1^2 + 6\alpha_1\alpha_2 + 2\alpha_1\alpha_3 + 3\alpha_2 \\
 &\quad + \alpha_3) - q^2(24\alpha_1^5 + 12\alpha_1^4 - 24\alpha_1^3\alpha_2 - 48\alpha_1^3 - 48\alpha_1^2\alpha_2 \\
 &\quad + 8\alpha_1^2\alpha_3 + 12\alpha_1^2 - 6\alpha_1\alpha_2^2 + 30\alpha_1\alpha_2 + 12\alpha_1\alpha_3 + 3\alpha_2^2 \\
 &\quad + 18\alpha_2 + 6\alpha_3) + q(24\alpha_1^5 - 12\alpha_1^4 - 36\alpha_1^3\alpha_2 + 4\alpha_1^3\alpha_3 \\
 &\quad - 12\alpha_1^3 - 6\alpha_1^2\alpha_2^2 - 12\alpha_1^2\alpha_2 + 14\alpha_1^2\alpha_3 - 9\alpha_1\alpha_2^2 + 30\alpha_1\alpha_2 \\
 &\quad + 12\alpha_1\alpha_3 - 3\alpha_2^2) + \alpha_1(12\alpha_1^2\alpha_2 - 4\alpha_1^2\alpha_3 + 6\alpha_1\alpha_2^2 \\
 &\quad - 12\alpha_1\alpha_2 - 6\alpha_1\alpha_3 + 3\alpha_2^2)
 \end{aligned} \tag{67}$$

$$\begin{aligned}
 C_2 &= 2q^3(12\alpha_1^4 - 18\alpha_1^3 - 12\alpha_1^2\alpha_2 + 6\alpha_1^2 + 6\alpha_1\alpha_2 + 2\alpha_1\alpha_3 + 3\alpha_2 \\
 &\quad + \alpha_3) - q^2(24\alpha_1^5 + 12\alpha_1^4 - 24\alpha_1^3\alpha_2 - 48\alpha_1^3 - 48\alpha_1^2\alpha_2 \\
 &\quad + 8\alpha_1^2\alpha_3 + 12\alpha_1^2 - 6\alpha_1\alpha_2^2 + 30\alpha_1\alpha_2 + 12\alpha_1\alpha_3 + 3\alpha_2^2 \\
 &\quad + 18\alpha_2 + 6\alpha_3) + q(24\alpha_1^5 - 12\alpha_1^4 - 36\alpha_1^3\alpha_2 + 4\alpha_1^3\alpha_3 \\
 &\quad - 12\alpha_1^3 - 6\alpha_1^2\alpha_2^2 - 12\alpha_1^2\alpha_2 + 14\alpha_1^2\alpha_3 - 9\alpha_1\alpha_2^2 + 30\alpha_1\alpha_2 \\
 &\quad + 12\alpha_1\alpha_3 - 3\alpha_2^2) + \alpha_1(12\alpha_1^2\alpha_2 - 4\alpha_1^2\alpha_3 + 6\alpha_1\alpha_2^2 \\
 &\quad - 12\alpha_1\alpha_2 - 6\alpha_1\alpha_3 + 3\alpha_2^2)
 \end{aligned} \tag{68}$$

$$\begin{aligned}
 C_3 = & -q^2(12\alpha_1^5 - 12\alpha_1^4 - 12\alpha_1^3\alpha_2 - 6\alpha_1^2\alpha_2 + 2\alpha_1^2\alpha_3 + 6\alpha_1\alpha_2 \\
 & + 2\alpha_1\alpha_3 + 3\alpha_2^2 + 6\alpha_2 + 2\alpha_3) + q\alpha_1(12\alpha_1^4 - 12\alpha_1^3 \\
 & - 18\alpha_1^2\alpha_2 + 2\alpha_1^2\alpha_3 - 3\alpha_1\alpha_2^2 + 4\alpha_1\alpha_3 + 12\alpha_2 + 4\alpha_3) \\
 & + \alpha_1^2(6\alpha_1\alpha_2 - 2\alpha_1\alpha_3 + 3\alpha_2^2 - 6\alpha_2 - 2\alpha_3).
 \end{aligned} \tag{69}$$

These expressions were factored by the computer using the "divide if divisible" subroutine and a long list of trial divisors including all factors appearing in (59). The complete solution of (64) was obtained in three passes through the computer with some assistance from the author between passes. On the first pass the quantity

$$B_1\Phi_{01} + C_1$$

was calculated [using (61) for Φ_{01}] and an attempt was made to divide it by each of the factors of A_{13} . The division by the cubic factor was successful, and there followed the result

$$\Phi_{02} = \frac{2qF - G}{6(q - \alpha_1)^2} \tag{70}$$

where

$$\begin{aligned}
 F &= 6\alpha_1^3 - 6\alpha_1^2 - 6\alpha_1\alpha_2 + 3\alpha_2 + \alpha_3 \\
 G &= 12\alpha_1^3 - 12\alpha_1^2 - 6\alpha_1\alpha_2 + 2\alpha_1\alpha_3 - 3\alpha_2^2.
 \end{aligned} \tag{71}$$

On the second pass the quantity

$$A_{23}\Phi_{02} + B_2\Phi_{01} + C_2$$

was calculated [using (70) for Φ_{02} and (61) for Φ_{01}] and an attempt was made to divide it by each of the factors of A_{22} and by the numerator of Φ_{02} . The latter division was successful, and the result

$$\Phi_{11} = \frac{[2qF - G][q^2 - q(\alpha_1 + 3) + \alpha_1]}{12(q - \alpha_1)^2[q^2 - q(\alpha_1 + 2) + \alpha_1]} \tag{72}$$

was thereby obtained. On the third pass the quantity

$$[q^2 - q(\alpha_1 + 2) + \alpha_1][A_{32}\Phi_{11} + A_{33}\Phi_{02} + B_3\Phi_{01} + C_3]$$

was calculated (the first factor being introduced in order to cancel the corresponding factor in the denominator of Φ_{11}) and an attempt was made to divide by each of the factors of A_{31} . Only the division by q was successful. This yielded the final result

$$\beta_2 = \Phi_{20} = \frac{(2qF - G)(q^2 - 2q)}{6(q - \alpha_1)^2[q^2 - q(\alpha_1 + 2) + \alpha_1]} \tag{73}$$

where F and G are the expressions defined in (71). Note that Φ_{02} appears as a factor in Φ_{11} and Φ_{20} . In the absence of feedback ($p = 0, q = 1$) (73) reduces to

$$\beta_2^{(0)} = \frac{\alpha_2}{1 - \alpha_1} + \frac{\alpha_2^2}{2(1 - \alpha_1)^2} + \frac{\alpha_3}{3(1 - \alpha_1)} \quad (74)$$

which is the correct result. To see this we use the law of composition to obtain

$$\beta_2^{(0)} = \alpha_2 + 2\alpha_1\omega_1 + \omega_2 \quad (75)$$

where α_1 and α_2 are the service moments while ω_1 and ω_2 are the waiting moments. The latter are given in Ref. 5 as

$$\begin{aligned} \omega_1 &= \frac{\alpha_2}{2(1 - \alpha_1)} \\ \omega_2 &= \frac{\alpha_2^2}{2(1 - \alpha_1)^2} + \frac{\alpha_3}{3(1 - \alpha_1)}. \end{aligned} \quad (76)$$

Substituting (76) into (75) we obtain (74) as expected. To obtain (36), substitute (71) into (73), replace each α_k by $\lambda^k \alpha_k$, and divide the resulting expression by λ^2 .

REFERENCES

1. Takács, Lajos, *Introduction to the Theory of Queues*, New York, Oxford University Press, 1962.
2. Takács, Lajos, A Single-Server Queue with Poisson Input, *Operations Research* **10**, 1962, pp. 388-394.
3. Takács, Lajos, Delay Distributions for One Line with Poisson Input, *General Holding Times and Various Orders of Service*, B.S.T.J., this issue, p. 487.
4. Brown, W. S., A System for the Symbolic Manipulation of Polynomials in Several Variables on the IBM 7090 Computer, to be published.
5. Riordan, John, *Stochastic Service Systems*, New York, John Wiley and Sons, Inc., 1962, p. 50.

Contributors to this Issue

WILLIAM E. ADAMS, B.S.I.E., University of Florida, 1949; Southern Bell Telephone Co., 1949-1953; Bell Telephone Laboratories, 1953—. Mr. Adams first worked with the cable methods development group. He has since worked with carrier systems development and is presently with the transmission circuits group for L-carrier terminals. Presently attending Graduate School at Northeastern University, evenings. Registered Professional Engineer, Massachusetts.

W. G. ALBERT, Bell Telephone Laboratories, 1951—. Since his association with the Laboratories, Mr. Albert has worked on equipment development engineering for carrier telephone terminals and carrier supplies for broadband coaxial and microwave transmission systems. Member, IEEE.

E. G. ANDREWS, B.A., 1922, William Jewell College; W. E. Co., 1922-25; Bell Telephone Laboratories, 1925-1959. At Bell Laboratories Mr. Andrews was first concerned with engineering and maintenance specifications for central office equipment. During World War II he was engaged in the development of radar training devices and relay digital computers, and later was concerned with various phases of planning and programming for computers used in military systems. After retiring from Bell Telephone Laboratories in 1959, Mr. Andrews joined Sanders Associates, Inc., where he is presently Manager of Preliminary Design. Member, IEEE, American Association for Computing Machinery, American Ordnance Association; Professional Engineer, State of New Hampshire.

F. R. BIES, B.S.E.E., Cooper Union, 1930; Bell Telephone Laboratories, 1925—. Mr. Bies has worked on the development of quartz crystals and electrical filters for long and short-haul carrier systems. He is presently concerned with the development of filters and equalizers for broadband carrier terminal equipment.

OMER P. CLARK, B.S.E.E., 1941, State University of Iowa; Western Electric Co., 1941-43; Bell Telephone Laboratories, 1943—. Mr. Clark's

assignments at the Laboratories have included circuit design for a wide variety of radar equipment, both airborne and ground-based, and design of circuits for the digital computer for the Nike Zeus system. He is presently working on primary frequency supplies and transmission amplifiers for broadband carrier equipment. Senior member, IEEE, member Eta Kappa Nu.

THEODORE V. CRATER, B.S.E.E., 1947, Montana State College; M.S.E.E., 1949, California Institute of Technology; Ph.D., 1953, Northwestern University; faculty, Michigan College of Mining and Technology, 1949-50; Bell Telephone Laboratories 1953—. Mr. Crater has been engaged in studies of problems in pulse transmission over telephone cable pairs. At present, he is in charge of a group studying experimental video communication. Member, IEEE, Eta Kappa Nu.

HOWARD CRAVIS, M.A., Harvard University, 1949; Bell Telephone Laboratories 1953—; Communications Development Training Fellowship, 1956-1960. He has made transmission systems studies related to speech interpolation, video communication, and short-haul trunk carrier systems. Member, IEEE.

A. DESCLOUX, Math. Dipl., École Polytechnique Fédérale (Swiss Federal Institute of Technology), 1948; Ph.D., Mathematical Statistics, University of North Carolina, 1961. After spending 1955-56 on the staff of the University of Washington where he taught mathematics and statistics, Dr. Descloux joined Bell Telephone Laboratories in 1956. At the Laboratories, he has been concerned chiefly with the application of probability theory to traffic problems. Member, Institute of Mathematical Statistics, American Mathematical Society, and Society for Industrial and Applied Mathematics.

ELBERT J. DRAZY, B.S. in E.E., 1942, Purdue University; Bell Telephone Laboratories, 1942—. During World War II he was concerned with the development of test equipment for microwave radar. Between 1951 and 1962 he was engaged in the development of test equipment for microwave radio relay systems and video transmission systems and of carrier supply equipment for the L-type multiplex. At present he is concerned with the development of FM terminals for the TD3 radio relay system. Member, Eta Kappa Nu, Tau Beta Pi, Sigma Xi.

JAMES B. EVANS, JR., Sc.B., Brown University, 1947; M.S.E.E.,

Worcester Polytechnic Institute, 1949; Bell Telephone Laboratories, 1949—. Mr. Evans' first assignment was the development of filters for coaxial carrier systems; he later worked on the design of thermistors and on development of short-haul carrier systems. He is presently supervisor of a group working on development of carrier frequency supplies for frequency-division multiplex terminals.

JOHN J. GINTY, B.S., 1950, and M.S., 1951, Boston College; M.B.A., 1962, Northeastern University; Bell Telephone Laboratories, 1956—. Mr. Ginty was first engaged in improvement program work for long-haul coaxial carrier systems. More recently he has been involved in the design and development of solid-state circuits for carrier and carrier terminal systems.

R. SHIELDS GRAHAM, B.S.E.E., University of Pennsylvania, 1937; Bell Telephone Laboratories, 1937—. Mr. Graham did graduate work at Columbia University and Polytechnic Institute of Brooklyn. He first worked with the design of equalizers and filters for use in broadband transmission systems. During World War II, he designed circuits for electronic fire control computers. Since then he has headed groups working on network design, computer analysis of systems, pulse transmission systems, multiplex terminals, and presently heads a group working on transistor feedback amplifiers. Mr. Graham is a senior member of IEEE and a member of Tau Beta Pi and Pi Mu Epsilon.

HERMANN K. GUMMEL, Dipl. Phys., 1952, Philipps University (Germany); M.S., 1952, and Ph.D., 1957, Syracuse University; Bell Telephone Laboratories, 1956—. His work has been in research and development of semiconductor devices. Member, American Physical Society, Sigma Xi.

F. J. HALLENBECK, E.E., 1936, Polytechnic Institute of Brooklyn; Western Electric Co., 1923-25; Bell Telephone Laboratories, 1925—. For many years he was involved in the development of transmission networks for Bell System and military communication facilities. He later supervised a group engaged in the development of broadband carrier systems. In 1958 he assumed responsibility for L-carrier terminal development. Senior member, IEEE; member, Tau Beta Pi, Eta Kappa Nu.

JOHN B. HARLEY, B.S.E.E., Brooklyn Polytechnic Institute, 1936; Western Electric Co., 1924-25; Bell Telephone Laboratories, 1925-26, 1928—. Mr. Harley's early work was in the field of carrier transmission

of radio broadcasts and the design of audio amplifiers and radio receivers. He later worked on sound recording and on systems development of mobile radiotelephone equipment. More recently he has been engaged in systems development of broadband microwave transmission facilities and is presently engaged in the design of carrier supplies for multiplex systems. Senior member, IEEE, member, Eta Kappa Nu and Tau Beta Pi.

J. J. MAHONEY, JR., A.T.&T.Co., 1926-34; Bell Telephone Laboratories, 1934—. At the Laboratories, Mr. Mahoney has worked on studies of protection systems for outside plant, the design of cathode-ray oscilloscopes used in radar test equipment, and systems engineering studies for carrier and wideband data transmission systems. He presently supervises a group responsible for transmission engineering of long-haul carrier systems. Member, IEEE.

DIETRICH MARCUSE, Diplom. Vorpruefung, 1952, and Dipl. Phys., 1954, Berlin Free University; D.E.E., 1962, Technische Hochschule, Karlsruhe, Germany; Siemens and Halske (Germany), 1954-1957; Bell Telephone Laboratories, 1957—. At Siemens and Halske Mr. Marcuse was engaged in transmission research, studying coaxial cable and circular waveguide transmission. At Bell Laboratories he has been engaged in studies of circular electric waveguides and work on gaseous masers. Member, IEEE.

R. E. POWERS, Bell Telephone Laboratories, 1953—. Mr. Powers' early work at the Laboratories was in the systems engineering area, where he was concerned with telephone signaling problems. In 1955, he was assigned to transmission development at the Merrimack Valley Laboratory, where, initially, he participated in developments concerning short-haul carrier telephone systems. Later, he was associated with the development of a wideband data transmission system. At present, he has responsibility for a number of the systems aspects of the L Multiplex development, including the regulator circuits described in the present article. He is attending the Graduate School at Northeastern University, evenings. Member, IEEE.

W. ROSENZWEIG, B.S., 1950, Rutgers University; M.S., 1952, University of Rochester; Ph.D., 1960, Columbia University. Brookhaven National Laboratory, 1951-53; Radiological Research Laboratory, Columbia University, 1953-1960; Bell Telephone Laboratories, 1960—. At Bell Laboratories, Mr. Rosenzweig has been mainly engaged in

studies of radiation damage to semiconductors. Member, American Physical Society, Radiation Research Society, Sigma Xi, Phi Beta Kappa.

IRWIN W. SANDBERG, B.E.E., 1955, M.E.E., 1956, and D.E.E., 1958, Polytechnic Institute of Brooklyn; Bell Telephone Laboratories, 1958—. He has been concerned with analysis of military systems, particularly radar systems, and with synthesis and analysis of active and time-varying networks. Member, IEEE, Eta Kappa Nu, Sigma Xi, Tau Beta Pi.

FRIEDOLF M. SMITS, Dipl. Phys., 1950, Dr. rer. nat., 1950, University of Freiburg, Germany; research associate, Physikalisches Institut, University of Freiburg, 1950-54; Bell Telephone Laboratories, 1954-62. Mr. Smits went to the Sandia Corporation in May 1962. His work at Bell Telephone Laboratories included studies of solid-state diffusion in germanium and silicon, device feasibility, and process studies, as well as the development of UHF semiconductor devices. He supervised a group that conducted radiation damage studies on components, particularly solar cells, used in the Telstar experimental satellite. Member of the American Physical Society and the German Physical Society.

LAJOS F. TAKÁCS, Doctor's Degree, 1948, University of Technical and Economical Sciences, Budapest; Doctor of Mathematical Sciences, 1957, Hungarian Academy of Sciences; Tungsram Research Laboratory (Telecommunications Research Institute), Budapest, 1945-55; Research Institute for Mathematics of the Hungarian Academy of Sciences, 1950-58; Roland Eötvös University, Budapest, 1953-58; Columbia University, 1959—; consultant, Bell Telephone Laboratories, 1959—. At present he is teaching probability theory and stochastic processes, and he is engaged in research in the mathematical theory of telephone traffic. Fellow, Institute of Mathematical Statistics. Member, American Mathematical Society, Mathematical Association of America, Society for Industrial and Applied Mathematics, American Statistical Association, Sigma Xi.

D. C. WELLER, A.B., 1946, Kenyon College; M.S.E.E., 1948, University of Illinois; Bell Telephone Laboratories, 1948—. His early work at the Laboratories included development of telephone carrier systems and exploratory development of a PCM carrier system. More recently he has engaged in computer development and is presently head of a group working on design of portions of the broadband carrier terminal equipment. Member, IEEE and Eta Kappa Nu.

