The Bell System Technical Journal

January, 1930

Telephone Communication System of the United States[1]

By BANCROFT GHERARDI and F. B. JEWETT

This paper presents the results which have been obtained up to the present time in developing telephone communication in the United States of America, this development having been worked out in a form to meet the particular conditions which present themselves in that country. The paper first deals with a brief description of the general structure and organization of the telephone communication system giving the organization of the Bell System which handles the greater part of the telephone service of the country and the reasons for and advantages of this organization. In this connection some figures are presented with respect to the technical personnel who are continuously engaged in studies to develop the art and to provide new methods and facilities for improving the service.

Local service, that is the service within the limits of a single telephone exchange area, is next discussed. Figures are given with respect to the volumes of telephone calls handled in the Bell System, the speed with which the connections for these calls are completed and the operating force required. Reference is also made to the standards of transmission given and the various problems encountered in meeting these standards. Figures are given with respect to station growth, to the increased efficiency of station apparatus and to the improvement in types of instruments. Various types of private branch exchanges provided to meet the needs of customers using a large amount of telephone service are discussed. The cable plant is considered mainly from the construction standpoint and typical illustrations are given of some of the construction practices. The various types of central office switching systems in common use are described, including magneto, common battery and dial systems, the latter including both the step-by-step and panel systems which are being provided in increasing amounts in the Bell System. The subject of buildings to house these various equipments as well as the operating forces and headquarters staffs in many cases is briefly discussed, also standardized layouts and floor plans. The problem of giving telephone service in the rural communities, which is a very important one in the telephone development in the United States, is also briefly treated.

The toll service is considered, first with respect to the shorter haul toll business and the problems involved and then with respect to the long distance toll service. Figures are given showing the speed of service and the amount of traffic handled. For the short distance toll service, two important methods of handling the business are described, namely, manual straightforward tandem and dial tandem.

The long distance service, which has developed most rapidly in recent years, is described in some detail in the paper. Among the important features of this service is noted the recently developed method of completing toll calls with sufficient speed so that on most of the calls the calling subscriber remains at the telephone. The various types of toll circuits are described including open wire circuits operated both at voice frequencies and by carrier systems and long toll cable circuits. The operation of these long circuits requires a large number of repeaters in tandem and the design and maintenance problems which this arrangement requires are pointed out in the paper.

[1] Presented by Dr. F. B. Jewett before the World Engineering Congress, Tokio Japan, October, 1929.

Information is given with respect to international telephone connections in North America, between North America and Europe and other international connections. In covering this subject some of the important items relating to the operation of the transatlantic radio channels are given and reference made to the projected transatlantic telephone cable.

Various forms of special services closely allied with the message telephone service are described. These include telegraph service, telephone circuits provided for private use, foreign exchange service, telephone networks for program transmission to radio broadcasting stations, electrical transmission of pictures, telephony in connection with aircraft operation, ship to shore telephony, telephony to mobile stations such as railroad trains, telephone services of railroads and other public utilities, telephone public address systems and television. Reference is also made to some of the by-products of the telephone development work which include improvements in submarine cable telegraphy brought about by the discovery of the alloys known as "permalloy and perminvar," the development work in the reproduction of sound and in the talking motion pictures.

In concluding, the paper points out that careful studies of the future development of the telephone industry indicate a somewhat accelerated rate of development of the services required to meet the demands of the customers and a continuing very rapid technical development of telephone plant and systems to provide the necessary facilities.

In treating such a large subject in a paper of this kind it has been necessary to deal with technical problems in rather general terms and as an attachment to the paper references are made to numerous articles in the technical press for the more technical information.

GENERAL

THE purpose of this paper is to give a general description of the telephone communication system of the United States of America, outlining briefly some of the more important engineering problems involved and indicating the service results obtained. At the beginning of this paper it seems important to give a brief description of the general structure and organization of the telephone communication system.

The commercial telephone system of the United States is entirely owned and operated by corporations, partnerships, and individuals. A group of 24 closely associated Bell Telephone Operating Companies owns and operates 14.8 million telephones and the telephone lines used for toll service within their territories. In addition there are in the country about 4.7 million telephones owned by several thousand independent telephone companies which have operating agreements with the Bell Companies providing for the interconnection of lines, thus permitting the operation of 19.5 million telephones as a single system. There are in addition about 140,000 telephones in the country not connected with the Bell System.

The 24 Bell Operating Companies cover the entire area of the United States and are responsible for all Bell Telephone operations within their respective areas. A number of the larger companies are subdivided into autonomous operating units, there being at the present time a total of 34 such units in the country. In many cases the area within corporate limits or within the limits of an operating unit is identical with that of a major political subdivision of the United States (a State) and this simplifies the application of governmental regulation. A typical organization of a Bell Operating Company is indicated in Fig. 1.

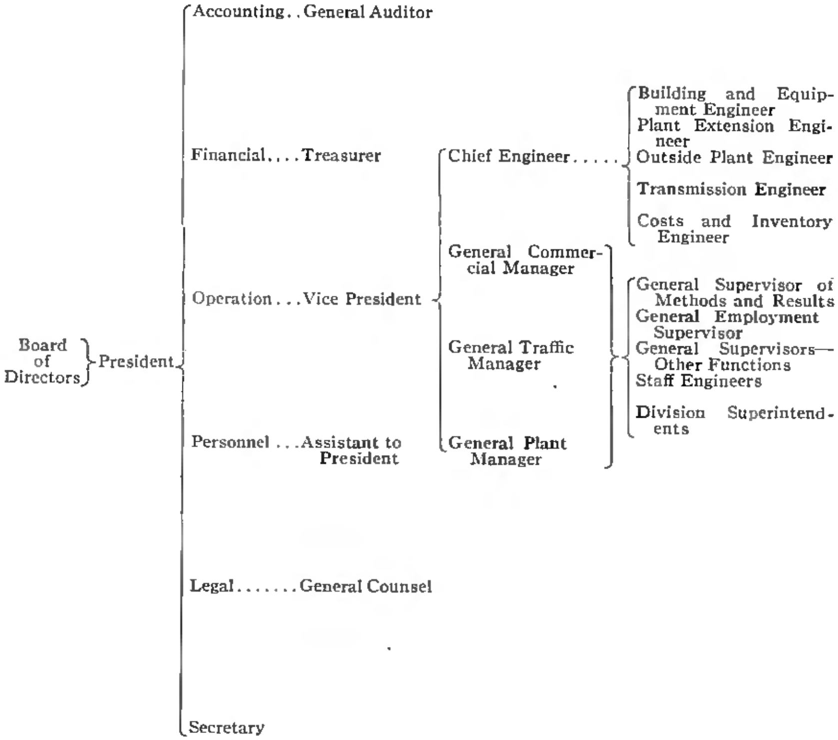


Fig. 1—Organization of typical Bell Telephone Operating Company.

In order to facilitate the best possible handling of the long distance service between points in different operating companies and to avoid the problems which would arise from divided responsibility, the long distance business involving territories of two or more Associated Companies is handled throughout the country by the Long Lines Department of the American Telephone and Telegraph Company. These operations are, of course, in the closest cooperation with the operations of the Associated Companies without duplication of construction or of operating effort.

An important feature of the Bell Telephone System is the general departments maintained by the American Telephone and Telegraph Company, including the Bell Telephone Laboratories. These departments, constituting about 7,500 engineers, scientists, business experts and assistants, are continuously engaged in studies to develop the art and to provide methods and facilities for improving the service. They also provide consulting advice to the operating companies on all phases of the telephone business and render to them a large variety of services. One of these services is making available to all the companies rights under all patents necessary for the fullest and most economical development of the business. It is the intention, in general, that work which can best be done once for all the entire telephone system rather than individually by the several operating companies shall be done by these general departments and that the specific solution of the telephone problems in each area shall be the responsibility of the operating company involved, who, however, are free at all times to get the advice and assistance of the general staff. In all of the work of the general staff close contact is maintained with the various operating telephone companies of the Bell System. The experiences of these companies are studied and analyzed to make available for all the companies the valuable results to be derived in this way, and the advantages to be obtained by comparing the experiences of different companies under similar conditions.

The organization of the American Telephone and Telegraph Company and the Bell Telephone Laboratories is indicated in Fig. 2.

Another very important feature of the Bell System is the very close relation between operating and manufacturing branches of the work through the ownership by the American Telephone and Telegraph Company of the Western Electric Company, Inc., and arrangements between that company and the operating companies for the supply of telephone apparatus and materials. This permits the manufacture of apparatus and the purchase of materials from outside suppliers to be done on the basis of the large quantities required for the entire Bell System resulting in great economies.

The organization of the Bell Telephone System is such as to result in close cooperation between the companies dealing with different branches of telephone work. This brings about the conditions necessary for universal service, for the development of the art along orderly and non-conflicting lines, and for the standardization of all apparatus, communication systems and operating methods to the extent that such standardization is helpful.

New types of telephone plant, operating methods, methods of maintenance and business methods are standardized by the general departments of the American Telephone and Telegraph Company, and are adopted and placed in use by all of the Associated Operating Companies to the extent that they apply to their local conditions. Special arrangements are, of course, made available to meet special requirements. The specifications for all standardized apparatus and materials are prepared by the Bell Telephone Laboratories, and the Western Electric Company is enabled to concentrate on the task of purchasing and manufacturing standardized supplies, materials and apparatus in accordance with these standard specifications. Standardization also has great operating advantages in minimizing stocks

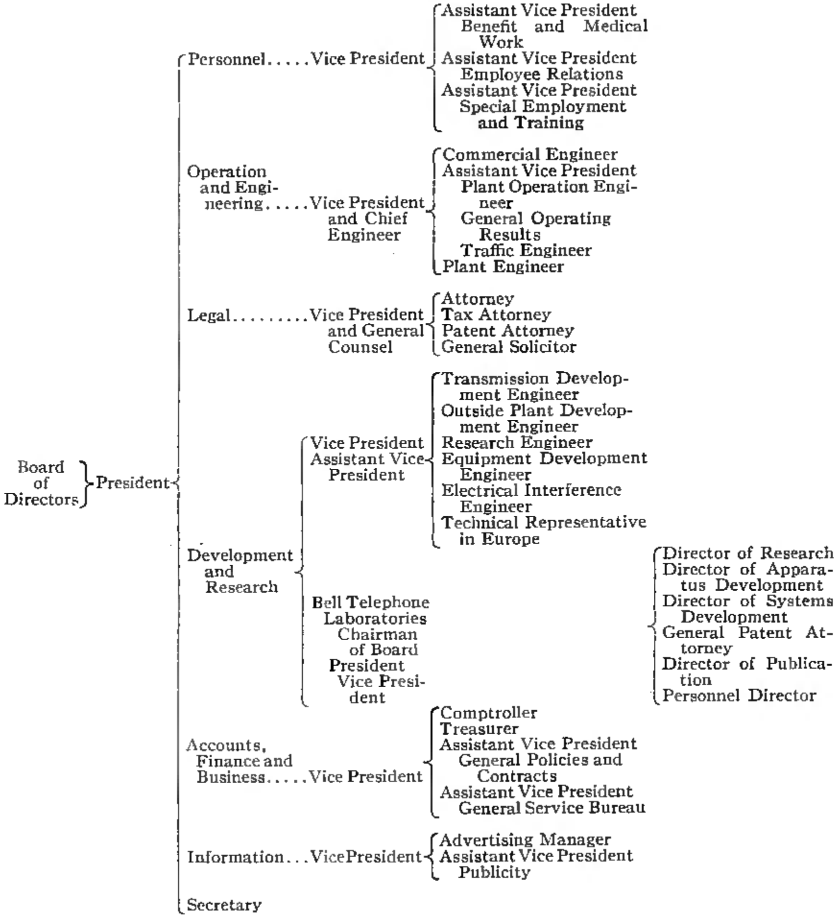


Fig. 2—Organization of the general departments of the American Telephone and Telegraph Company, including Bell Telephone Laboratories.

of materials and providing interchangeability both of materials and working forces and is very important in making possible the operation of the entire interconnected network as a single system with uniform grades of service.

As has been stated by Mr. Gifford, President of the American Telephone and Telegraph Company, "The ideal and aim today of the American Telephone and Telegraph Company and its Associated Companies is a telephone service for the nation, free, so far as humanly possible, from imperfections, errors, or delays, and enabling at all times any one anywhere to pick up a telephone and talk to any one else anywhere else, clearly, quickly and at a reasonable cost." With this aim in view, continuous effort is made further to improve and to extend the service within the nation and also the telephonic connections to other nations. It is recognized also that changes in business and social conditions bring about repeated changes in the services desired by the people of the nation and in the character and appearance of facilities furnished to them. These facts, in addition to the onward march of the application of science, form an important basis for the continued study by the general staff of the development of all phases of the telephone system.

A few figures relative to the size and growth of the Bell System are helpful in an understanding of the more specific telephone problems which are discussed below. Such figures are included in the statistical summary appended to this paper and include data regarding telephone messages, numbers of telephones, miles of wire and amount of telephone plant.

In accordance with the general organization of the Bell System, the engineering problems involved in the design, construction and maintenance of the plant of each operating telephone company are the responsibility of the engineering department of that company. General studies of methods of improvement of service and the development of new apparatus and systems of communication, together with consulting engineering advice, are provided by the general departments.

For the provision of new plant to meet additional demands for service, in the case of the more important items, often one year, and sometimes more, is required between the completion of detailed engineering plans and completion of construction. Furthermore, to obtain maximum economy it is necessary that much of the new construction provide for expected increases in demands for service for a number of years to come. This applies particularly to telephone buildings and to runs of underground conduit and to a lesser extent to cables, pole lines and many other very important parts of the telephone plant. The engineering of the additions to the Bell Telephone System, now aggregating over 500 million dollars a year is, therefore, necessarily based on careful forecasts of the amount and type of business to be expected for a number of years in the future and good engineering judgment must be applied in determining the types, quantities and design of plant. These must take into account not only the expected amount of service required but also expected future changes in the character and standards of service demanded and in the apparatus and materials expected to become available. In view of the capital expended in extensions and the large amount of plant already in service, the engineering work involved is considerable. There are now approximately 10,000 engineers engaged in the work of the Bell System of which approximately 6,300 are in the operating companies, 2,200 in the headquarters departments and 1,500 in the Western Electric Company. These figures apply to men doing work of engineering grade, and inclusion of assistants of all kinds, stenographical, clerical, laboratory, etc., would more than double these figures.

Local Service

General

Service within the limits of a single telephone exchange is spoken of as local service. This generally includes service within a large metropolitan area, a city with its surrounding suburbs or a town or village. During 1928 customers of the Bell System originated approximately 24,000 million local calls of which approximately 19,000 million originated from manual and 5,000 million from dial telephones. This represents an average daily usage of approximately 5.5 calls per telephone station per day.

The speed of service is illustrated by the following average figures. In the smaller cities with manual operation where the operator who takes the call completes it herself without trunking, the average time from the start of the call to the answer of the called station is 19 seconds. The corresponding figure for manual calls in large cities based on about three million observations made in the year 1928 in 38 large cities of the country is 28.8 seconds. The same observations indicate that when fully converted to the dial system the speed of service in the large cities will be about 22.5 seconds.

As to the accuracy of service, 98 per cent of all calls are handled without error. The most serious errors are those resulting in wrong numbers. The mistakes made by the subscribers and equipment under the dial system are about the same in number as those made by subscribers and operators under the manual system.

Calls resulting in busy reports amount to 10 per cent. This is something which is not directly under the control of the telephone company since the subscriber determines the telephone facilities which are provided. Records are kept, however, in both manual and dial offices of the lines responsible for the greatest number of busy reports and efforts are made to have the subscribers take additional facilities.

Standards of transmission are applied to the design of the plant to insure that transmission will be clear between the most remote parts of the exchange area. This depends on the design of station equipment, wire lines and switchboard equipment, and is expressed in terms of the combined electric and acoustic efficiencies of the circuits from the mouth of the talker to the ear of the listener. This overall efficiency is expressed in terms of the adjustment of a standard reference circuit. The standards in use in the United States refer to the maximum transmission loss permitted between any two subscribers and vary in magnitude between equivalents of 18 decibels and 22 decibels, depending on the circumstances of different cases.

In order to meet these transmission standards the Bell Companies have standard requirements regarding the efficiency of transmitters and receivers and other station equipment, and these are made the basis for engineering the wire plant. Transmission losses in switchboards are kept as low as practicable and within specified limits. The wire plant for subscribers lines and trunks is designed to be within the limits required for meeting the transmission standards. If under special conditions it appears desirable to exceed these limits, this is done only with the approval of responsible engineering authorities.

To handle calls at the local switchboards there was in the Bell System in 1928 an average operating force of about 122,000 young women. In addition an average force of approximately 36,000 were employed at the toll boards of the Bell System. This made a total operating force of 158,000. In order to make up for losses and for growth, 75,000 women were employed, and to select this number approximately 300,000 applicants were interviewed.

One of the important administrative problems is the scheduling of the operating forces so that an adequate number may be available in each central office throughout every period of the day. A method has been worked out whereby all types of operating work are equated to a common unit of measurements and the number of such units that an operator should handle to give the best service most efficiently has been determined. Frequent counts are maintained of the number of calls handled throughout each hour of the day and in this way the forces are adjusted to the work to be done.

In order that the demand for telephone service may be met promptly as it develops and further that plant additions may be along sound and economic lines, calls for careful planning. To this end the fundamental plans prepared for the different exchange areas forecast the telephone development from 15 to 25 years in the future. Such fundamental plans show the proposed central office locations, the boundaries of the districts to be served by each office, and the plan of the underground conduit system. They are based on analysis of the existing market for telephone service; the forecasted market at a future date, considering both growth and distribution of population; expected changes in wage levels; estimates of the amount of service that will be sold under probable future rate conditions; and other factors.

Station Apparatus

One of the most important parts of the telephone plant is the apparatus installed on the subscribers' premises known as the station apparatus. Of this equipment the telephone transmitter and telephone receiver are fundamentally important elements and continued research work has been carried out to improve the efficiency, clarity of reproduction and reliability of these instruments. As a result of improvements in transmitters, receivers and induction coils the overall efficiency, for example, of the station apparatus has since 1912 increased by a factor of 6.5. At the present time commercial transmitters when fully energized by direct current, are capable of delivering electrical energy in the form of voice currents 200 times as great as the acoustic energy of the voice of the speaker by which the transmitter is actuated. For the most important part of the frequency range used in speech this ratio of output power to speech power is considerably greater. That is to say, the transmitter acts as a high ratio amplifier.

In the Bell System the type of station equipment most generally in use is the desk stand. As the result of extensive development work it has been possible to produce a hand set which has transmission characteristics equal to those of the desk stand equipped with the best instruments heretofore in use. The hand set development involved the solution of difficult problems, the principal of which were to prevent singing or distortion of quality on account of the rigid connection between receiver and transmitter and to make the transmitter efficient through the wide range of positions in which it is placed by the user.

The latest form of this instrument is shown in Fig. 3. In addition to the usual black finish, this telephone as well as the bell box and other station apparatus have recently been made available in five colors, statuary bronze, old brass, oxidized silver, ivory and French gray.

Practically all the service for business purposes is provided by individual lines or by private branch exchanges as discussed later.

Fig. 3—The latest form of hand set.

For residences, however, there is in the United States a large development of two-party and four-party lines. The two-party stations are provided with selective ringing so that each station is signaled only for its own telephone messages, and the four-party stations are provided in some places with selective ringing and in others with semi-selective ringing.

Party lines have furnished a satisfactory means of providing service to small users and have been an important factor in the development of new fields of service in residences.

To care for situations where something more than a single line with one or two telephones is needed, but where an inter-communicating system or private branch exchange is not justified, so-called wiring plans are used which provide various arrangements for associating the station equipment with the telephone lines. For the most part the customers' needs are satisfactorily met by one of the ten standard arrangements in general use. A specific example is that of two central office lines with two main and two extension stations. Calls to or from either telephone line may be made from any one of the four telephones. Answering at a main station provides privacy by cutting off all the other telephones.

Fig. 4—Telephone booth provided for public telephone stations.

There are in the United States a considerable number of extension stations. At the present time there are in the Bell System over 1.3

million of such stations. This number is rapidly increasing particularly for residence use as people appreciate further the advantages of having telephones in a number of convenient locations. The best residences are more and more being equipped to have telephones available in all parts of the house.

In order to make telephone service possible for those people whose sense of hearing is more or less deficient, special sets are installed. By means of a vacuum tube amplifier which the user can adjust, the receiving may be amplified so as to bring the range up to the point giving best results, this point depending on the degree of impairment of his hearing.

Public telephone stations constitute an important part of telephone development in the United States, there being at present more than 275,000 of such stations in service. Whereas residence and business service is largely given by contract, the customers contracting to pay a definite amount per month or a certain amount per call, a great many of the public pay stations are supplied with coin boxes by means of which the money is collected at the time the call is made.

These installations are also for the most part in booths to insure quiet and privacy. Fig. 4 shows a form of booth furnished by the telephone companies, provided with a seat and with lighting and ventilation as well as a convenient location for the necessary telephone directories.

Private Branch Exchange

For customers who use a large amount of telephone service, one of several types of private branch exchange is provided which not only permits distribution of the incoming calls to the particular station desired but also makes it possible for one extension to call another without going through the central office. In the Bell System

Private Branch Exchange Stations	2,740,000
Per Cent of Total Bell Stations	19.0
Private Branch Exchange Boards	
Cordless	53,300
Cord	60,900
Total	114,200
Private Branch Exchange Cord Positions	
Manual	68,600
Dial	1,700
Total	70,300
Private Branch Exchange Attendants	
Cord Board Attendants	75,000
Cordless Board Attendants	53,000
Total	128,000

Fig. 5—Private Branch Exchange Statistics for the Bell System as of Feb. 1, 1929.

there are 36 million telephone connections handled each day by about 128,000 private branch attendants. About 17 per cent of all local calls originate at these boards. Equipment of both the manual and dial type is installed, the latter being particularly adapted to extension-to-extension calling. Further data regarding private branch exchanges are given in Fig. 5.

The smallest manual installation is the cordless board illustrated in Fig. 6 where connections are established by means of keys, and the capacity is limited to three trunks and seven stations. A larger type, using cords for the completion of connections, is illustrated in Fig. 7 with a capacity for fifteen trunks and 200 stations.

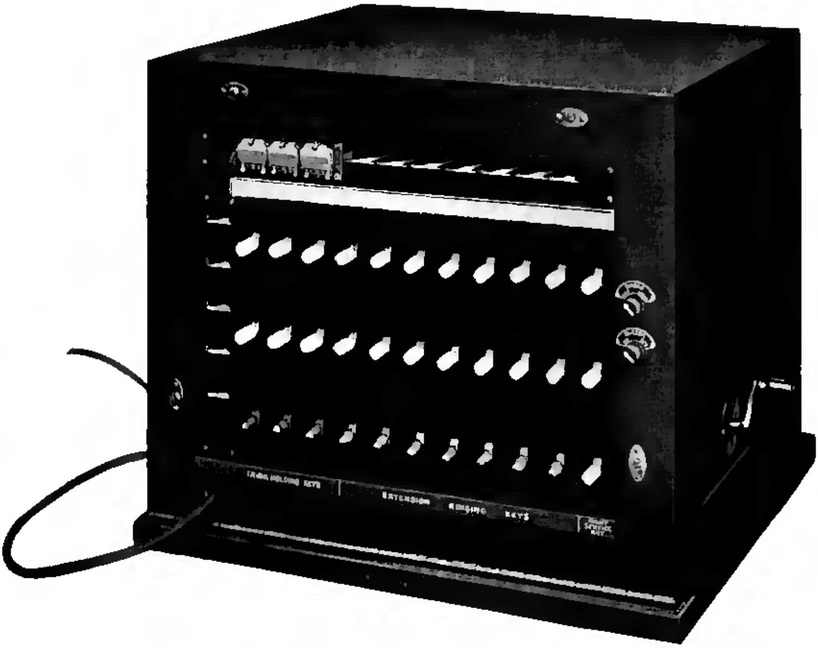


Fig. 6—Cordless private branch exchange installation for 7 stations and 3 central office trunks.

For the largest private branch exchanges the equipment is of much the same type as that used at central offices. A large switchboard for one of the public utilities having 1,600 stations is shown in Fig. 8. Connecting this private branch exchange with the central office there are 148 lines and in addition 151 tie trunk lines extend to other private branch exchanges having a business association. There are a total of 42 switchboard positions.

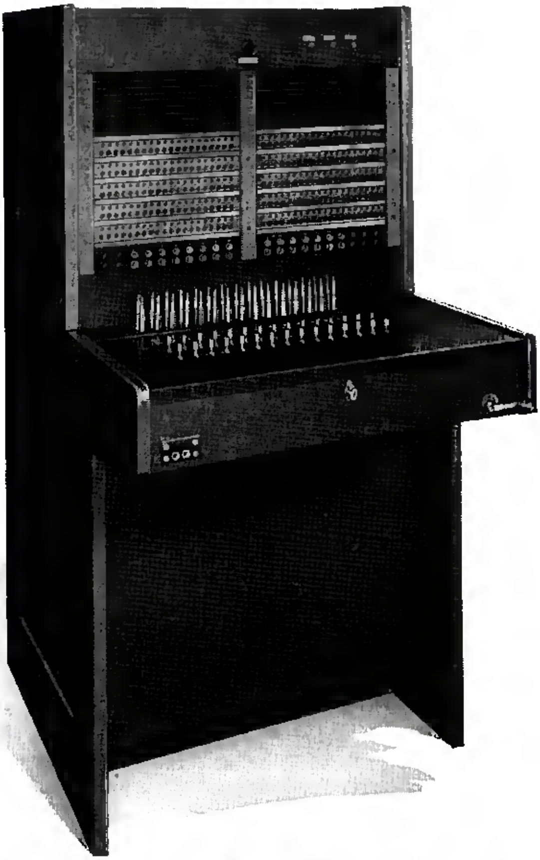


Fig. 7—Private branch exchange switchboard arranged for 15 trunks and 200 stations.

Fig. 8.—Private branch exchange—1600 stations, 148 lines to Central Office, 151 tie trunks to other private branch exchanges, 42 switchboard positions, 60 operators and 24 hour service.

At the end of 1928 in the Bell System there were about 650 dial P.B.X. installations with about 100,000 lines. The smaller sizes of this equipment are designed to meet the needs of the larger residences and the larger sizes are adequate for any business office and large industrial plant. Typical equipment arrangements are shown in Fig. 9.

In general the private branch attendants are in the employ of the subscribers having this type of equipment. It is essential that the attendants be recruited and trained with the same care as central office operators, and to this end, the telephone companies maintain employment bureaus and training courses for the benefit of private branch exchange attendants. Subscribers are encouraged to send their attendants to these training courses for retraining wherever this appears advisable.

Instructors highly trained in local and toll central office operation and in the best methods for handling private branch exchange work constantly visit private branch exchanges in order to be of assistance both to the subscribers and to the attendants in giving the best possible service.

Cable Plant

While open wires are occasionally used in limited quantity at outlying points, 96 per cent of the exchange area wire plant is in cable. Of this 74 per cent is underground.

An outstanding development is the steady increase in the number of pairs of conductors which it is possible to place in a single lead sheath of 6.7 cm. outside diameter. From the early use of 30 to 60 pair 19 gauge conductors there has been a continual increase in the number of pairs and a decrease in the size of the wires until at the present time 1,800 pairs of 26 gauge conductors are placed under a single sheath for use in the denser areas. The significance of this development is indicated in Fig. 10 which shows the year in which each important step was taken and the relative cost per pair of conductors resulting from each step in the development.

In urban development main cables, called feeder cables, usually of the maximum size, radiate from each central office through underground conduit to the various parts of the area served. These feeder cables in general are run full size for considerable distances rather than being diminished at branch cable points, and the flexibility thus obtained is found advantageous for conditions in this country. Each main feeder cable has smaller branch cables bridged to it at intervals. The main feeder cable may continue all the way through the area

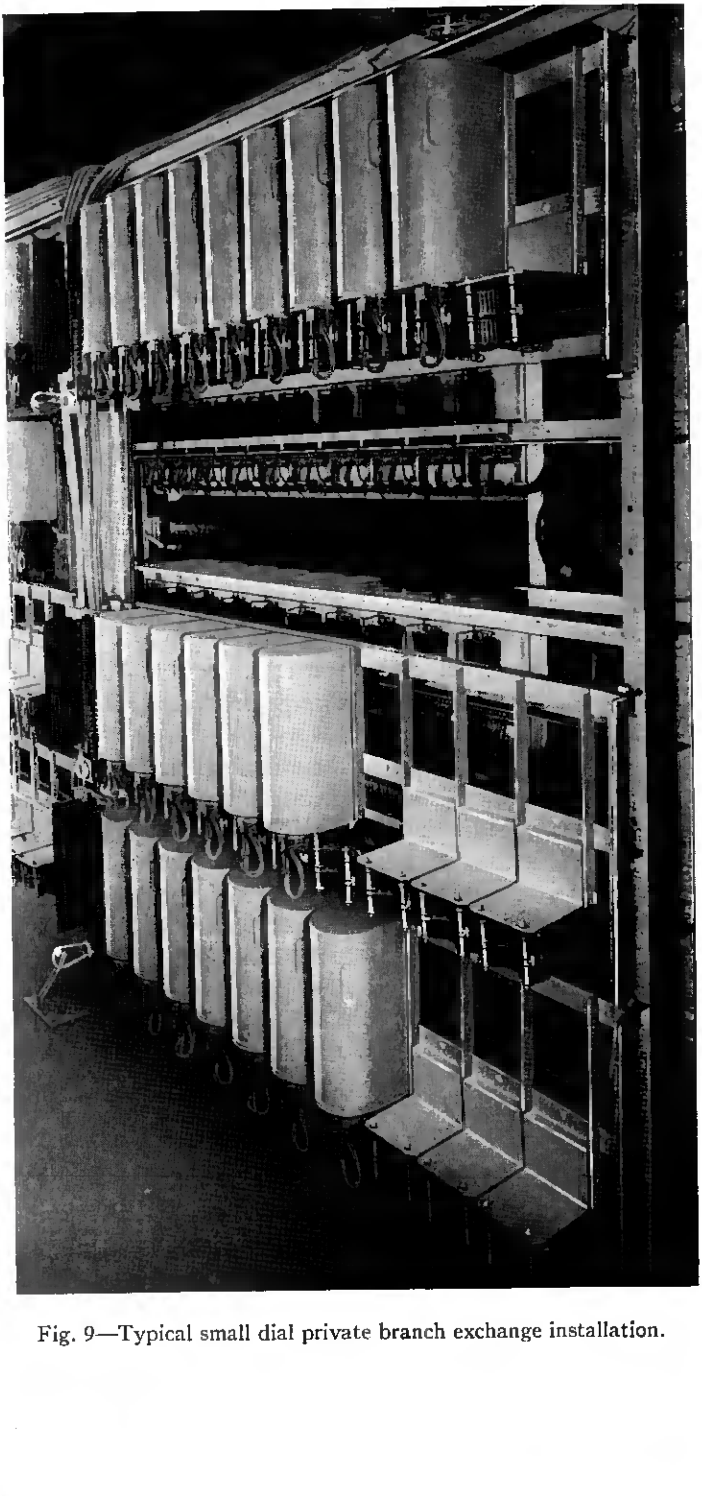


Fig. 9—Typical small dial private branch exchange installation.

or it may divide into two or more smaller cables branching out either underground or aerially in the area.

A type of distributing plant located along the rear properties of a residential block is illustrated in Fig. 11. This represents the usual

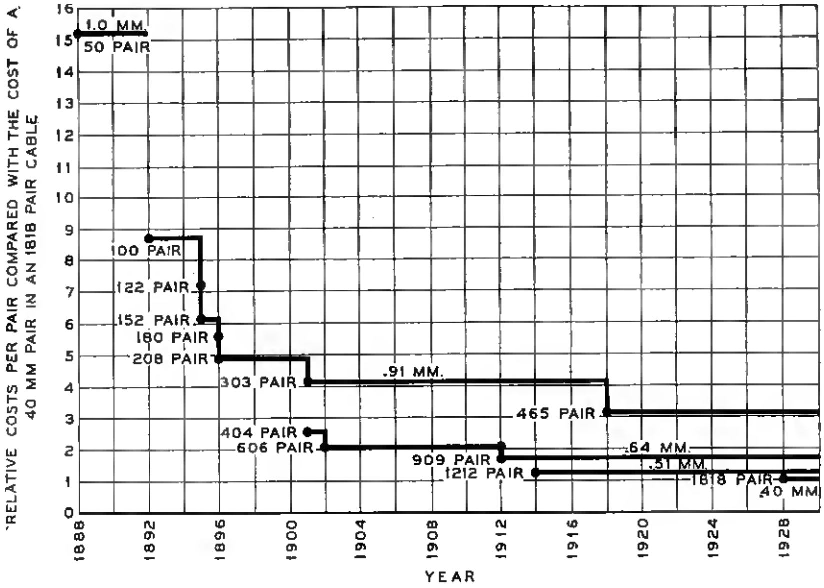


Fig. 10.

Fig. 11—Aerial cable distribution along rear property line. Poles used jointly for electric lighting power and telephone distribution.

type of construction followed in such areas. With continuous building construction the distributing cables are often attached to the rear walls as in Fig. 12 or extended through the basements of the buildings.

Underground cables are carried in conduit consisting of various combinations of multiple tile duct. A typical duct run, shown in Fig. 13, illustrates the materials and methods of construction generally employed.

Fig. 12—Distribution telephone cable attached to rear wall of building showing terminal boxes and entrance by twisted pair into cellar.

At the central office the conduit system is designed to meet the ultimate requirements of the building and terminates in a cable vault as shown in Fig. 14. With this entrance arrangement the main cables are spliced in the cable vault to smaller units of silk and cotton insulated conductors which extend up through the building in slots or ducts to the main distributing frame.

Switching Systems

The outstanding development in switching systems for a telephone communication has, of course, been the rapid trend toward an increase in the extent to which the operations are performed by automatic machinery. The general characteristics of the different types of switching systems in use in the United States and their extent of use are briefly discussed below.

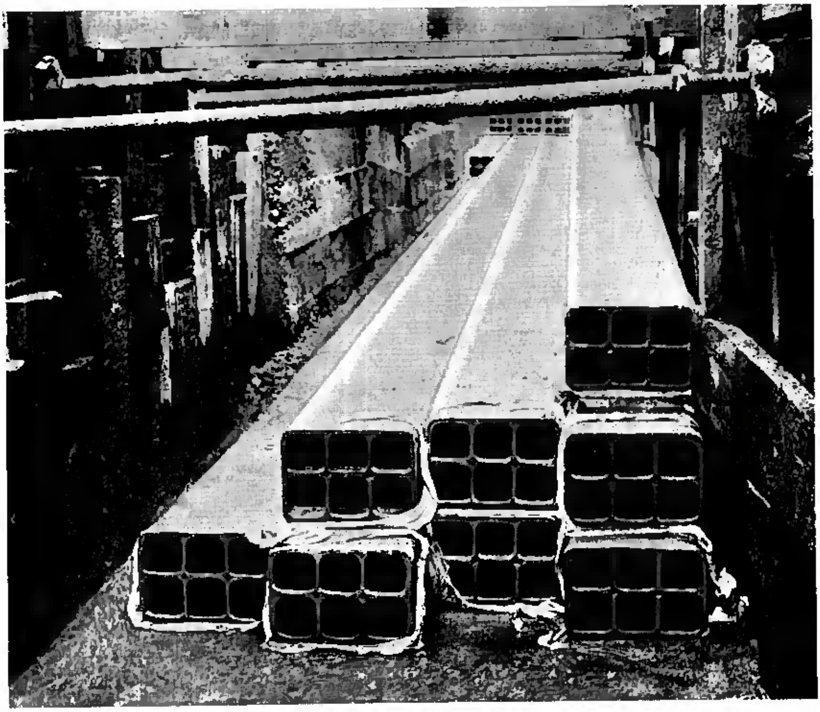


Fig. 13—72 duct underground cable run under construction.

Magneto

Magneto switching arrangements are used in small places and scattered rural areas. They vary in size from an arrangement to interconnect two or three lines up to a switchboard handling 300 to 400 lines. The average size of the magneto switchboards in the Bell

System is 170 lines. At the end of 1928 there were about 3,500 magneto offices with approximately 5,500 operators' positions, serving 1.1 million telephones.

Common Battery

At the end of 1928 there were 2,036 common battery offices, the maximum size office serving 10,500 lines and the average being 3,700 lines. There was a total of 46,000 switchboard positions where the subscribers' lines are answered. In addition, there were 13,000 so-called trunk positions which are required where it is necessary to trunk calls from one office to another in areas having more than one central office.

Fig. 14—Cable vault in Central Office, St. Louis, Missouri, showing entrance of cables through ducts and connecting to silk and cotton insulated cables extending up through the building.

The trend in development in manual switchboard has been toward performing more of the necessary switching and signaling operations automatically by means of somewhat more complicated circuit and equipment arrangements and less and less by the operator. These changes have resulted in less manual operating labor and in an improvement in the service. Some of the more important of these changes are as follows:

1. Automatic ringing which continues automatically at regular intervals until the subscriber answers.
2. Ringing tone, very much reduced in volume, to the calling subscriber automatically advising him when ringing is in progress.
3. The audible busy signal, a tone placed on the calling line when the called line is busy.

An important recent change in manual central office equipment relates to the trunking methods employed in completing a connection from one central office to another. In most of the larger cities the so-called "straightforward" method is used. With this operating plan the number that is desired in the distant office is passed by the originating or "A" operator to the completing or "B" operator over the trunk that is used for completing the trunk connection. This is in contrast to the "call circuit method" where all orders between operators are passed over a separate wire known as a call circuit. The principal equipment changes at the "B" positions have to do with the different circuit plans for connecting the trunk operator's telephone set to the trunk. This is done either by means of a key, by means of plugging the trunk into a listening jack or automatically by means of suitable relays. At the "A" end the principal change is the arrangement for testing whether or not an outgoing trunk is busy. This is done either by means of a lamp indicating a free trunk or by a lamp or tone indicating that a part of a trunk group is free.

Dial Equipment

Dial equipment of two types known as the step-by-step system and the panel system respectively are used in about equal amounts in the Bell System. The change from manual to dial operation presented a very large problem from an engineering, a manufacturing, an installing and an economic standpoint. At first the dial installations were to care largely for growth but they have been followed by installations for the replacement of existing manual equipment where, all factors considered, this was clearly justified. In this way an orderly program has been developed. Figure 15 indicates the total number of stations on a dial and on a manual basis for each year since 1921 and the expected program up to 1933. Under the present contemplated dial program it is estimated that the areas employing step-by-step equipment will be on a complete dial basis by 1937 and that all areas employing panel equipment will be substantially completed by 1942.

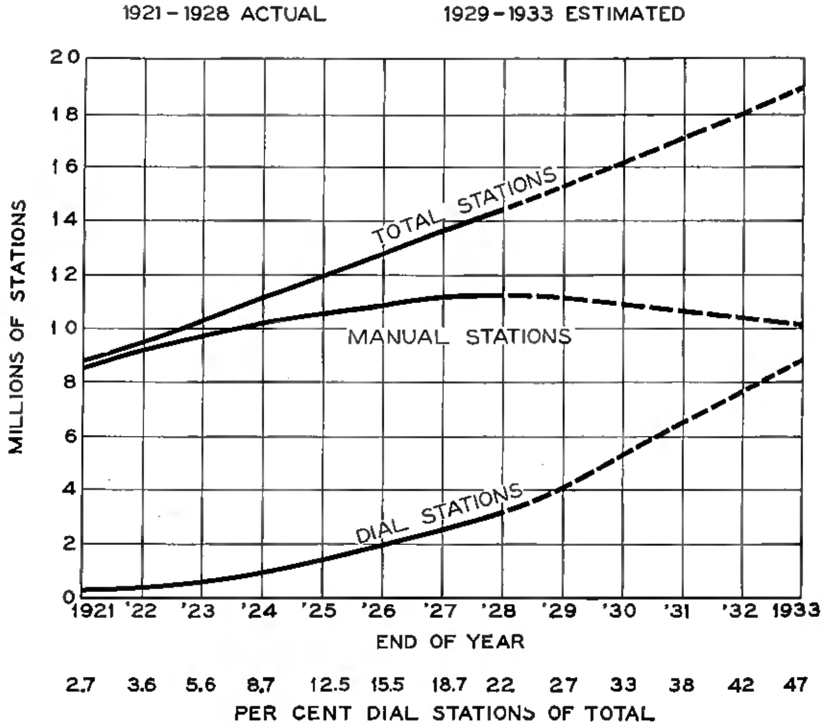


Fig. 15.

Fig. 15—Relation between manual, dial and total stations—Bell owned stations.

Step-by-Step Dial System

The step-by-step system is used in the Bell System in single office areas and in the smaller and medium sized multi-office areas where the number of central offices is limited and consequently the trunking problems are not complicated. Step-by-step equipment is in service in 194 offices to which are connected 1.6 million stations.

The fundamental unit of the step-by-step dial system is the selector illustrated in Fig. 16. This switch has a capacity for 100 terminals placed ten on a level and ten levels high, thus making possible the selection of any one of a hundred lines. By placing several selectors in series a network of central offices may be built up, each office serving 10,000 telephones.

The selectors are mounted on iron frameworks and the terminals are cabled to cross-connecting frames so that any grouping can be made as may be demanded by the number of calls which the particular selector handles. A typical arrangement of selectors is shown in Fig. 17.

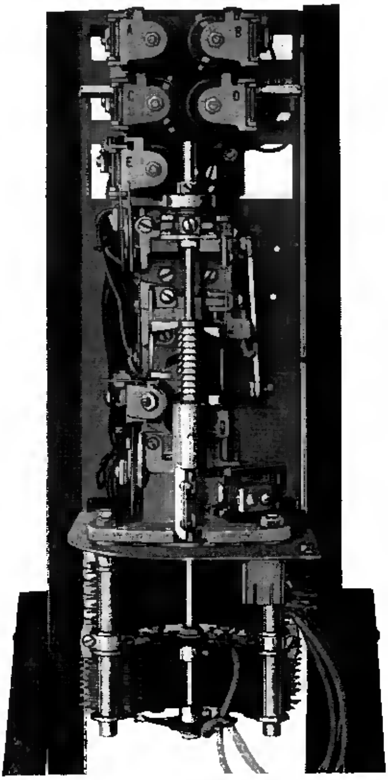


Fig. 16—Typical step-by-step selector showing relays which control the circuit operation mounted at the top, and the selector banks of 100 terminals at the bottom.

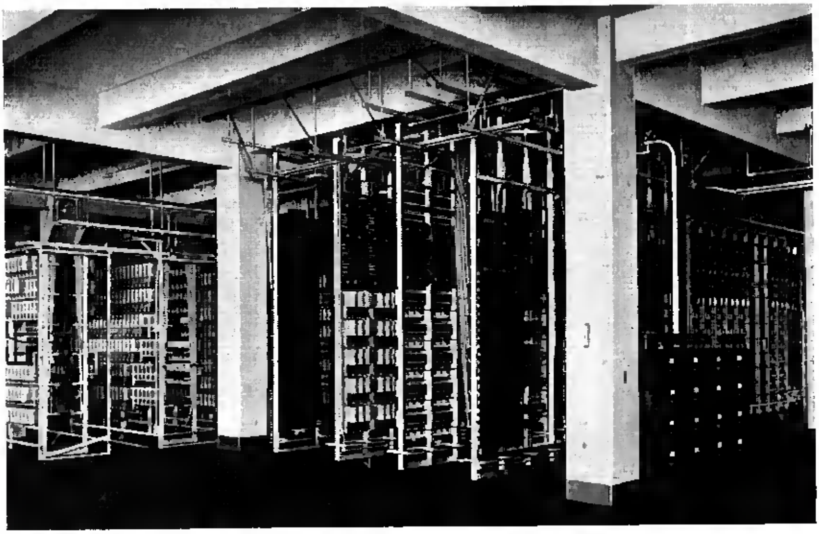


Fig. 17—Installation of step-by-step dial equipment showing selectors mounted on iron framework.

Panel Dial System

In the largest cities and the smaller municipalities around them making up the large metropolitan centers, a much greater degree of complexity is encountered in switching the calls due to the large number of offices of varying types to which calls are destined and due

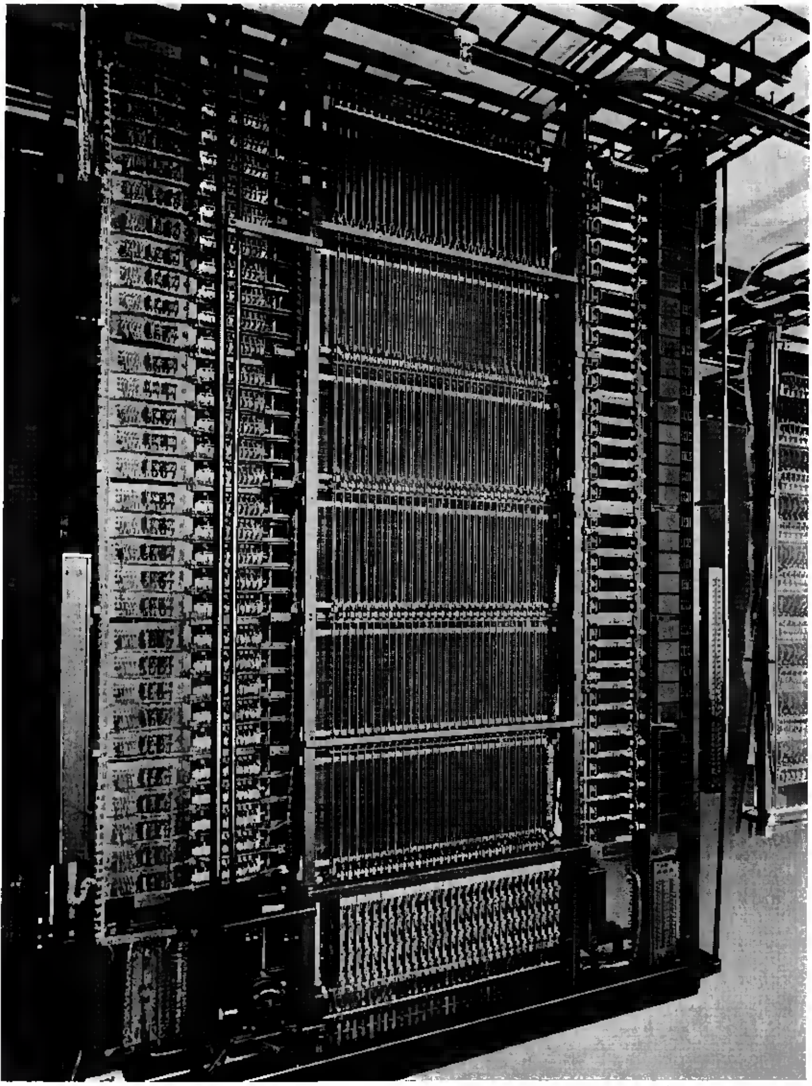


Fig. 18—Typical panel selector frame. Capacity 30 selectors in front and 30 selectors in rear. Motor driving mechanism is at the bottom of frame and controlling apparatus at either side.

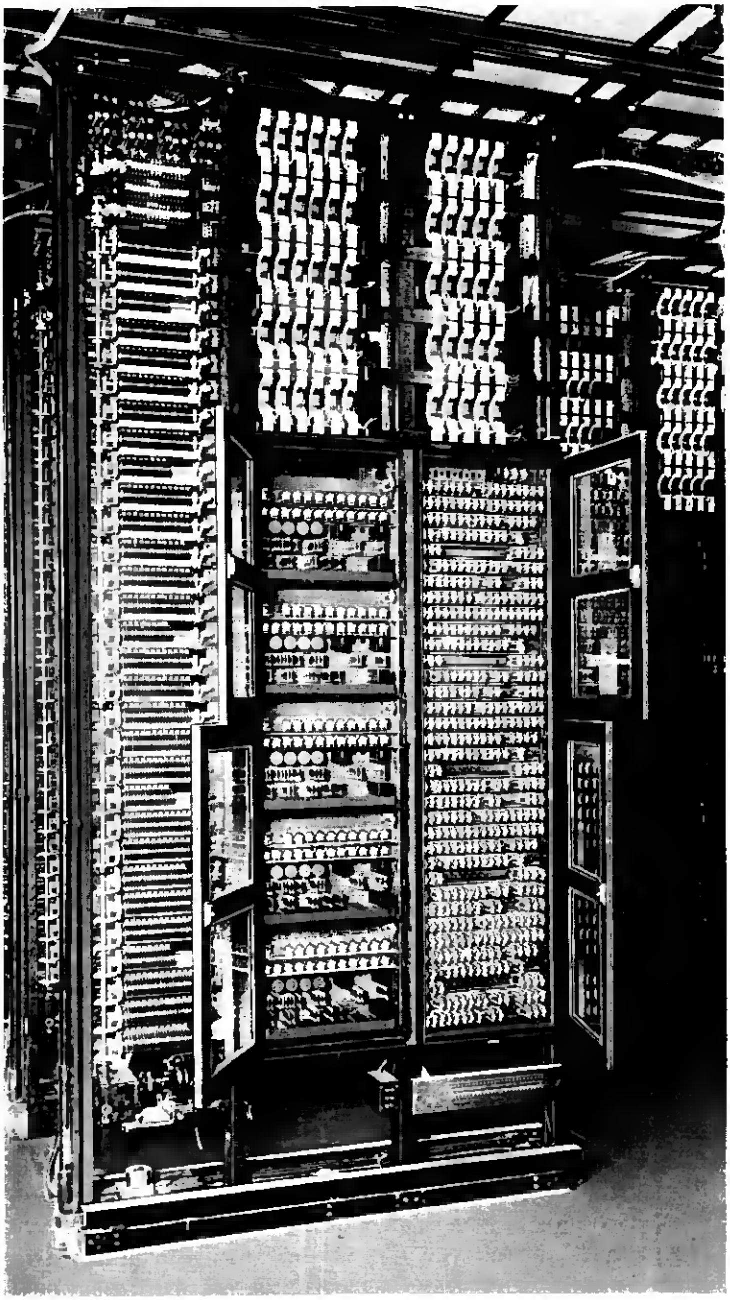


Fig. 19—Panel dial office sender frame showing the apparatus for five senders.

to the routings involved in order to trunk economically either large or small volumes of traffic. The panel system was developed to meet these requirements and is now installed in a number of such metropolitan centers, notably, New York, Chicago, Philadelphia, Boston, Detroit, Cleveland, St. Louis, Pittsburgh, Baltimore, San Francisco, Buffalo, Kansas City, Seattle, Providence and Omaha. Panel equipment is in service in 128 offices to which are connected 1.6 million stations.

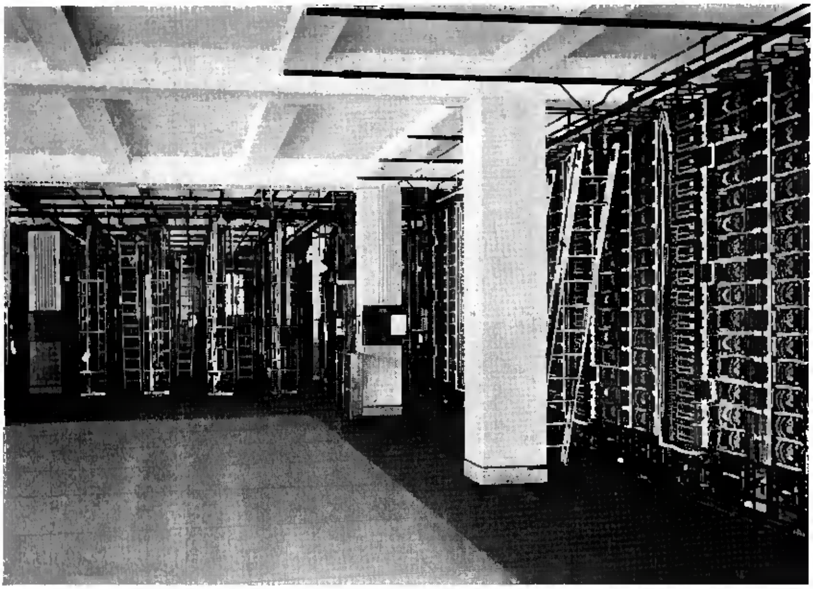


Fig. 20—Installation of panel dial equipment. The unused floor space is provided for future growth.

In the panel system the fundamental switching unit is a large switch consisting of five banks of 100 terminals each. The selectors, by which contact is made with any one of the 500 terminals, move vertically on both sides of the terminal banks. A typical panel frame having capacity for 60 selectors is illustrated in Fig. 18.

In the panel system the selectors do not follow in synchronism with the impulses of the dial as in the step-by-step system. Rather, a group of apparatus known as the "sender" records the impulses and in turn directs the operation of the several selectors in the train until the called terminal is reached. By this means the trunking arrangements and the numbering scheme can be designed independently of each other. This, combined with the large capacity of the panel selectors, makes possible economies in interoffice trunks and a reduction in the number of selectors involved in completing a connection in a large exchange. The selection may be either to a dial office or to a manual office reached direct or through a tandem switchboard. Figure 19 illustrates the apparatus for five senders. A switchroom in a typical panel office is shown in Fig. 20.

Buildings

The buildings in the Bell System at present number about 6,000, excluding those occupied by the Western Electric Company. All of the larger and many of the smaller are owned by the telephone companies. The range in size of the buildings is illustrated by Fig. 21 showing a small building for a single manual switchboard with a present capacity of 2,200 lines, and Fig. 22 showing a headquarters office and equipment building in a large city. This building has 66,000 square meters of floor space, in the lower 9 floors has space for dial equipment to serve 100,000 telephones, and in the upper floors includes offices for 5,000 people. Figures 23 and 24 further illustrate

Fig. 21—Telephone office in the residential district of Silver Spring, Maryland. Designed for small manual switchboard with a present capacity of 2,200 lines.

Fig. 22—Combined equipment and office building, New York City, containing headquarters of the New York Telephone Company, 31 stories, 66,000 square meters of floor space. Lower 9 floors arranged for toll tandem equipment and for dial equipment with an ultimate capacity of 100,000 telephones. Upper 22 floors arranged for offices with a capacity of about 5,000 people.

some of the recent combined equipment and office buildings for large cities.

The objective in connection with buildings of all types, including equipment and office buildings, garages and warehouses, is to provide

Fig. 23—Combined equipment and office building at Detroit, Michigan containing headquarters of the Michigan Bell Telephone Company. 19 stories, 28,000 square meters of floor space. 13 floors arranged for equipment including toll board and dial equipment to serve 60,000 telephones. 6 floors are arranged for offices.

Fig. 24—Combined equipment and office building at Cleveland, Ohio containing headquarters of the Ohio Bell Telephone Company. 22 stories, 25,000 square meters of floor space. 13 floors designed for equipment including the toll board and ultimate dial equipment for 100,000 telephones. 9 floors arranged for office space.

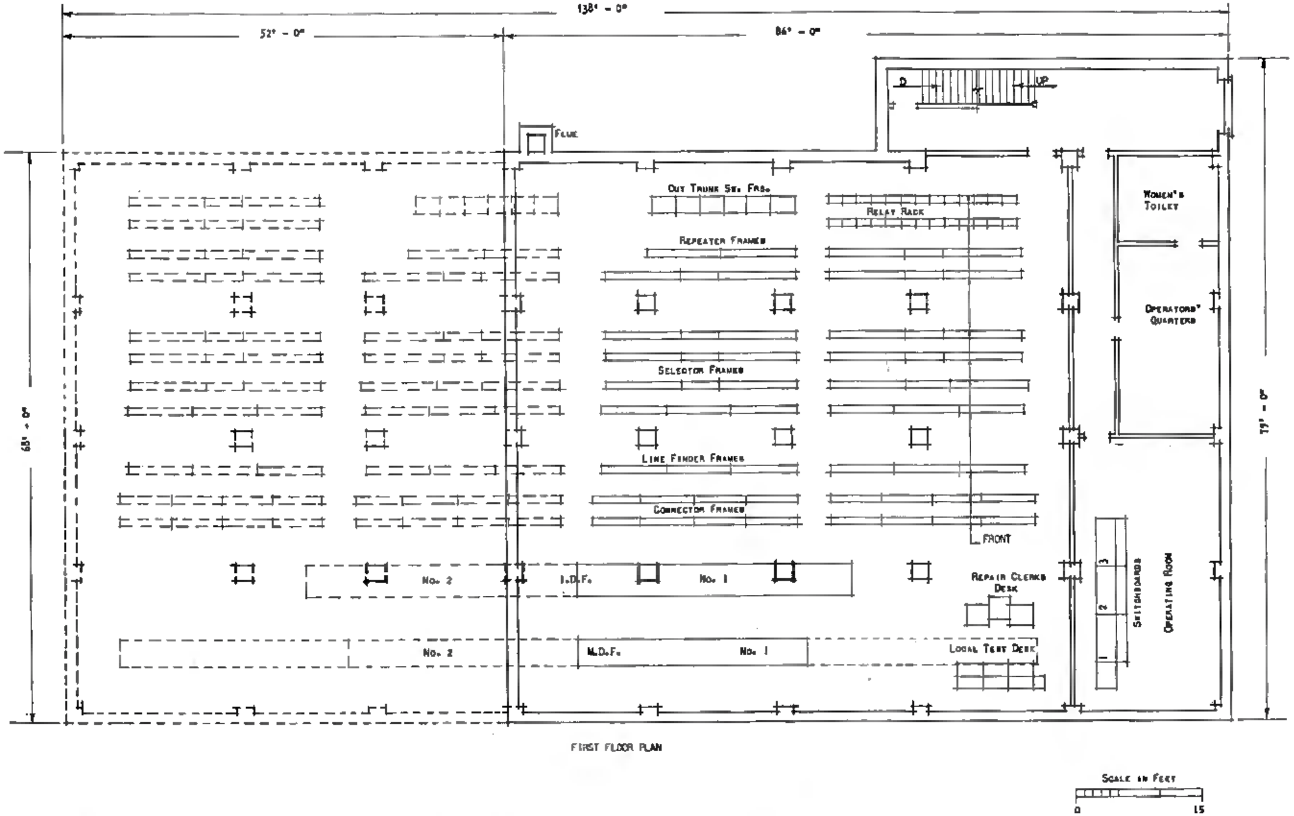


Fig. 25—Standard floor plan for step-by-step dial office designed for best arrangement including flexibility for future growth.

buildings which adequately, economically and comfortably house the equipment and personnel—both initially and throughout the useful life of the building—and which at the same time are outstanding and attractive, appropriate to their surroundings and a continuing source of satisfaction to the communities in which they are erected.

The initial size of a building is determined by the costs and rate of increase in space requirements, with due regard to the service reactions caused by extensions of the building which is being used for operating purposes. As a result of these considerations central office buildings are usually designed with a capacity of about twice the initial requirements. In many cases space provided for later extensions of equipment is used temporarily for office space. Possibilities of future extensions of the buildings are provided for either by buying more land than is required for the initial building, thus making possible lateral extensions, or by providing strength in the steel framework for future vertical extensions. The possibility and type of future extensions must, of course, also be taken account of in the architectural design.

Floor plans have been developed representing for typical conditions the best arrangements of the various types of central office equipment both manual and dial. The use of these uniform floor plans greatly facilitates the engineering, manufacture and installation of the equipment, and results in savings of both time and money. A uniform floor plan applying to a step-by-step dial office is illustrated in Fig. 25. The relative location of the different frames and aisle space as well as the unit size of the frames is fixed but the number of frames is varied to meet the requirements of individual cases.

Except for the smallest buildings, non-combustible construction is used, a steel or reinforced concrete frame and brick or stone curtain walls being employed. The very small buildings, except where severe fire exposures are encountered, are of frame or brick and joist construction.

Rural Service

Surrounding the larger cities and towns there is in general a sparsely settled district developed on a multi-party basis. Service is given on common battery lines where the distances are not too great and either semi-selective or code ringing is used. These lines usually serve not more than six or occasionally eight parties and the common battery signaling requirements limit the range to about seven or eight kilometers.

One of the most difficult service problems of the Bell System is that of providing service to farming districts where the distances between successive farms as developed in the United States is often great. At the present time service in such farming areas is usually provided by multi-party lines with magneto signaling. These rural lines carry from six to as many as fifteen parties and may be as much as sixty-five kilometers in length.

In the past the demand for service from rural districts of this nature has generally been limited almost wholly to local service between the rural customers and shorter haul toll service, and the present development is a response to that point of view.

The extent of development of rural service is illustrated by a census made in the State of Iowa in 1920, showing that of 213,000 farmers 86 per cent have telephone service. This percentage is doubtless materially higher at the present time. There were in 1928, in the Bell System, 6,000 offices which served lines that might be classified as rural. Of these rural lines about 12,000 were on a common battery basis and about 43,000 on a magneto basis. These figures do not include the rural lines which were served by connecting companies, the addition of which would increase the above figures many times. The development of this type of service has to a very considerable extent been in the hands of local groups of small local companies because of the nature of the service.

With the present rapid development in the use of a nation-wide toll service, there are a rapidly increasing number of stations where rural customers will accept a higher grade of service designed for general connection to telephones throughout the Bell System. The development of improved rural service of this type is an important feature of the present telephone program in the United States.

TOLL SERVICE

Service between telephones which are not in the same local exchange area is called "Toll Service." With the exception of less than one per cent of the telephones which are not connected to the Bell System, toll service is offered in the United States between any two telephones in the country and to a very large extent the toll plant is adequate to provide good service between any two of these telephones.

An outstanding feature of the last few years has been the rapid growth of the toll service. This is indicated in the appended statistical summary which shows that during the last five years the completed toll messages have increased by 67 per cent. During this same period the number of telephones in service have increased by about 28 per cent. These figures show that in spite of the continued increase in the number of telephones in service, the number of toll messages per telephone have increased by about 30 per cent for this period.

One important cause of the rapid increase in toll usage has been the material improvements in toll service.

Figure 26 shows the increase in the speed of toll service since 1920 expressed in terms of the average time required from the placing of a call to the response of the called party, or until the operator gives a definite report regarding the call. The service is sufficiently fast so that on 95 per cent of the calls, the subscriber stays at the telephone. This makes possible still more rapid service and simplified operation.

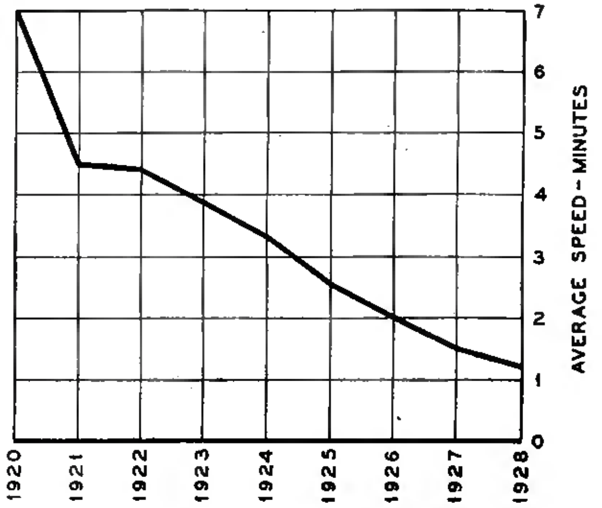


Fig. 26—The average time required from the placing of a toll call to the response of the called party, or until a definite report is made by the operator.

There have also been very great improvements during this period in the clearness of speech transmission. The maximum permissible transmission loss between two subscribers on a toll connection has been materially reduced. The toll plant and subscriber plant are now so designed that most of the messages are handled with a maximum transmission equivalent for the longest subscriber lines of 20 to 25 decibels overall referred to the standard transmission reference system.

SHORT DISTANCE TOLL SERVICE

General

To a large and increasing extent the toll messages are completed and supervised by the local exchange operators who first answer the subscriber's call, providing in this way toll service with the same methods which are applied to local service and with comparable speed. This method of operation is used for most of the toll business up to a distance of 50 kilometers and to a considerable extent up to 100 kilometers. The use of this method includes the extensive suburban areas around the large cities. Calls handled by this method now amount to 650 million messages a year and its increasing use has

42
4
32

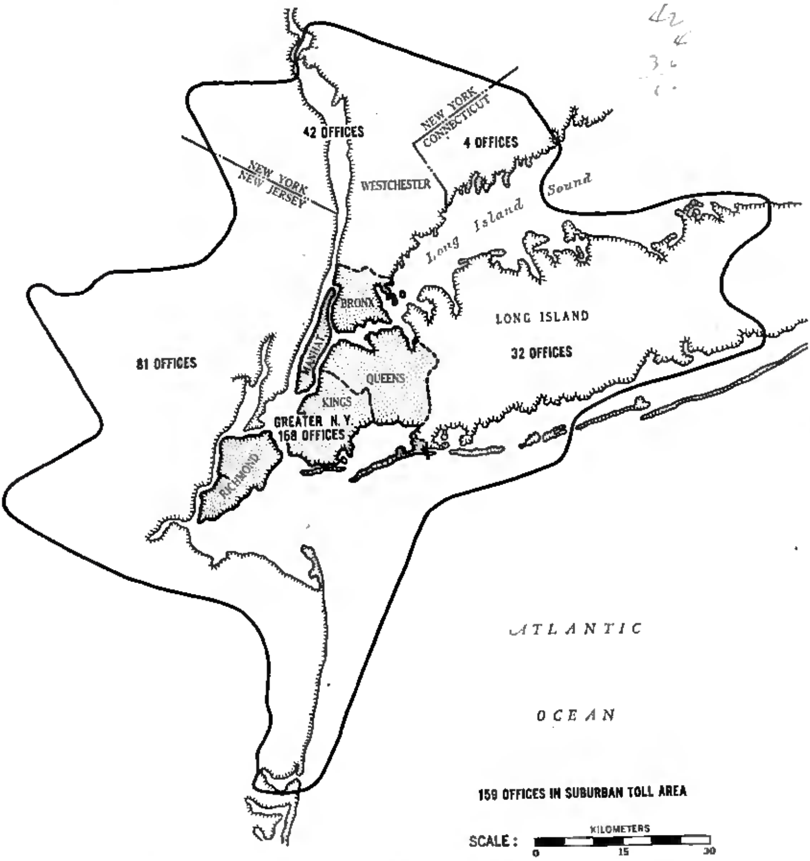


Fig. 27—New York suburban toll area indicating the number of offices in Metropolitan and suburban districts. Population in the area 10.5 millions. Telephones in the area 2.5 millions.

been one of the important ways in which increased speed of service has been brought about.

The handling of this suburban telephone traffic adds greatly to the complexity of the transmission, trunking and operating problem of the larger cities. This is illustrated by Fig. 27 showing the suburban toll area surrounding New York City. It will be noted that the city itself includes 168 central offices and in the suburban areas in the metropolitan district there are in addition 159 central offices.

In many cases of the shorter haul toll service, the volume of traffic between two offices is sufficient to warrant direct trunks and the calls are completed over these direct trunks by the usual local traffic methods. In order to provide an efficient trunking arrangement for the smaller volume of traffic between widely separated offices, however, tandem trunking arrangements are provided, by which the calls are routed through a central switching point and from that point distributed to the terminating offices. Either manual or dial central office equipment is used as outlined below, each type having its field of application depending upon the amount of traffic and the portion of traffic to and from manual and dial central offices. The trunks to the central switching point, or tandem office, are in general of somewhat larger gauge than interoffice trunks because of the greater distances involved and the correspondingly more severe transmission requirements. In some of the longer trunks telephone repeaters are used. It has not been found generally economical to use repeaters at the tandem switching point although this is done in certain instances and it is possible that in the near future the more general use of repeaters in this way may become an economical means of meeting the transmission requirements.

Manual Straightforward Tandem

The manual straightforward tandem is used in those medium sized areas in which most of the suburban calls are between manual switchboards. The arrangement of the equipment is shown in Fig. 28. The tandem trunks from the originating office terminate on the plugs located at the rear of the keyboard and the tandem completing trunks to the various terminating offices appear in jacks in the face of the switchboard. The completing trunks are provided with lamp signals indicating idle trunks. When a call comes in on a tandem trunk the operator is advised by a flashing lamp signal on that trunk and her telephone set is automatically connected to it. The work of the tandem operator is limited to making the connection at the tandem board and making the disconnection when advised by lamp signal that the conversation is over. Her work is thus greatly simplified and the operation of the board is extremely rapid.

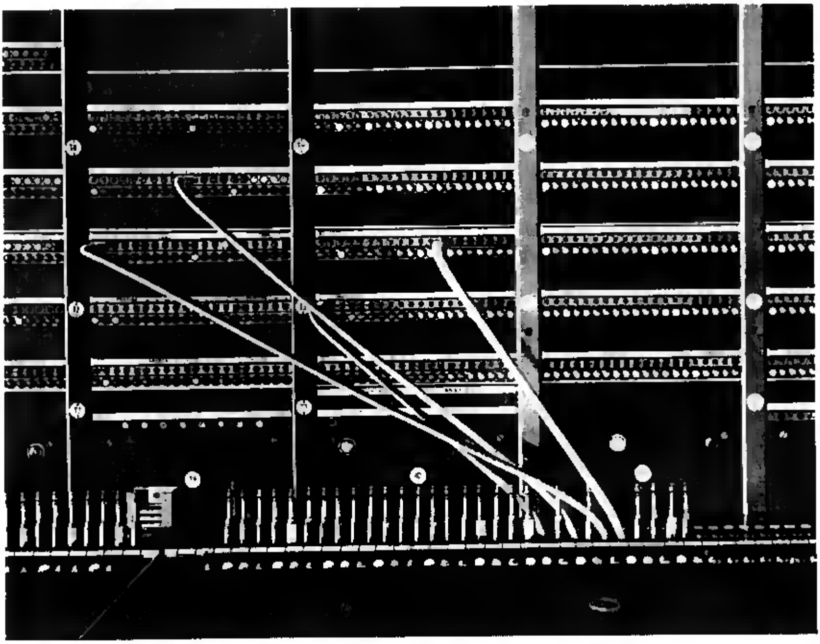


Fig. 28—Manual straightforward tandem board for handling suburban toll calls. Trunks from originating offices terminate on plugs at the rear of the keyboard. Tandem completing trunks appear in jacks in the face of the switchboard. Incoming calls indicated by flashing of lamp associated with trunk cord.

Dial Tandem

Dial tandem systems have been designed for handling suburban toll traffic in areas where a large proportion of the central offices are of the dial system and also to facilitate handling the complex suburban traffic around the large metropolitan areas, such as New York, Boston, Chicago and Philadelphia. The type of dial equipment, in general, corresponds with the type used for the local traffic in the same area.

For use in connection with calls from dial offices, the dial tandem usually requires no operators at the tandem office. The originating operator in the dial office who handles the short haul traffic controls the selection of the trunk to the terminating office.

In some of the large metropolitan centers, dial tandem apparatus of the panel type is employed also for handling calls from manual

Fig. 29—Panel type dial tandem office in New York City.

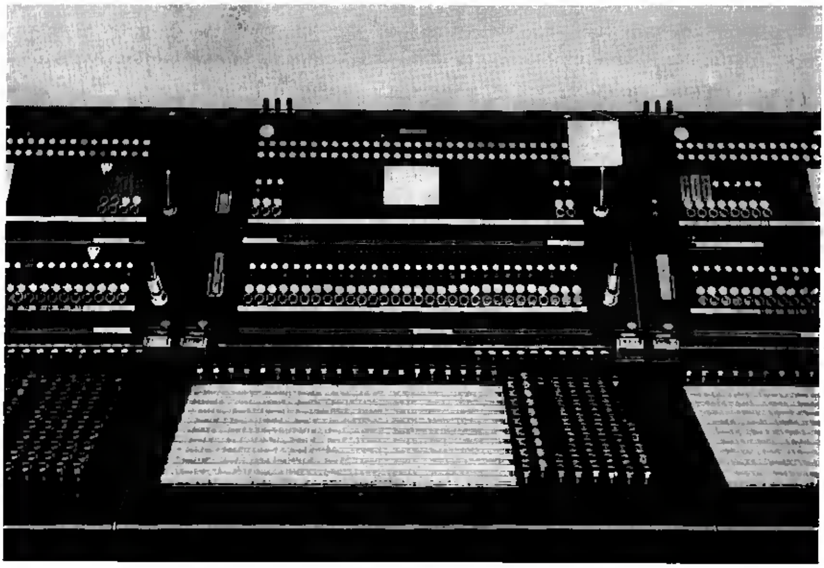


Fig. 30—Panel type dial tandem office. Arrangement of keyboard at tandem positions. Incoming trunks appear as lamps and associated keys on upper sloping part of keyshelf. Calls are completed by setting up the called office and called number on keys on lower part of keyshelf.

offices. With this arrangement operators are, of course, required at the tandem board. The tandem operator receives a request from the originating operator for the office and number called and by means of keys establishes the connection by dial switching apparatus to the called subscriber in case he is in a dial office or transmits the required information to the terminating office operator in the case of a manual office. Figure 29 shows an installation of panel tandem operators' equipment and Fig. 30 shows one section of this equipment in greater detail.

Fig. 31—Step-by-step dial tandem board. Calls completed to the terminating dial subscriber station by means of the 10 button key set shown on the keyboard. Incoming calls automatically distributed to an idle operator.

A modified form of step-by-step tandem equipment using operators has been installed in step-by-step areas for handling calls from manual offices to dial offices in cases in which it was not advisable to equip all the subscriber operators' positions with dials. This equipment includes, in addition to the selectors, a simplified type of switchboard as shown in Fig. 31.

Long Distance Service

General

For the longer hauls the subscriber is connected to a toll board operator who completes and supervises the toll message. This method is called the toll board method of operation, and is used for most all the messages over about 100 kilometers. For the purpose of this paper this service will be referred to as long distance service. The messages handled by this method total about 300 million messages a year. The amount of long distance business at New York City, for example, requires the use of 1,275 operators' switchboard positions.

During the past three years an important change has been generally applied in the methods of handling long distance service. Formerly the toll operator first receiving the call recorded the necessary information on a ticket and forwarded this ticket to another operator provided with facilities for completing calls to the particular part of the country involved in each case. An increase in speed has been brought about by providing the operators with arrangements both for recording calls and for completing calls to all points so as to avoid, in a large proportion of the cases, the necessity for transmitting the information to a second operator.

By means of this change in method and other improvements, the average speed of service for all long distance messages has been decreased from 6.9 minutes in 1925 to 2.6 minutes in 1928. Also in 1928, 90.7 per cent of the calls made by the customers resulted in completed messages.

In placing a long distance call, the telephone subscriber in the United States may give simply the telephone number and city desired. This has some advantages in speed of service. At present, 50 per cent of the long distance messages are handled in this way and this per cent is increasing. About 15 per cent of the messages are handled in this same way, the called telephone number, however, being supplied by the operator. In addition, the telephone system offers, for a somewhat greater charge, what is called a "particular person" service. This means that the subscriber may, if he wishes, ask to talk with a specified person at a distant point, giving such information as he can regarding how that person may be located. The telephone operator then undertakes to complete this message by locating the desired party, following him up to points other than that designated, if necessary, and if the calling subscriber wishes. The percentage of messages handled on this basis increases with the length of haul.

When a subscriber wishes he may transmit to the telephone company in advance information regarding a number of calls which he wishes to have completed in sequence, beginning immediately or at a specified time. These sequence calls, as they are termed, are used particularly in connection with selling by long distance telephone. At the present time at the New York long distance office, for example, about five per cent of the business is in the form of sequence calls, some of the sequence lists including as many as 1,000 calls.

Except during times of emergency conditions, it is not the practice in the United States to limit the length of conversations on toll connections. As a result it is very general for conversations, particularly on longer hauls, to exceed the initial three-minute period, the average length for transcontinental conversations being, for example, six

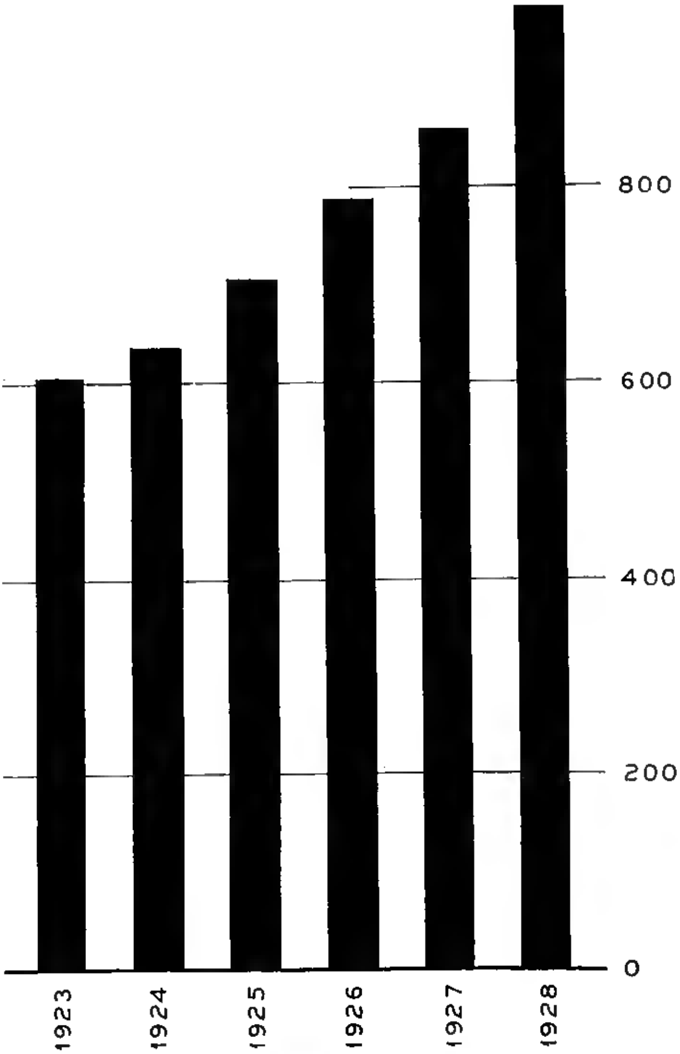


Fig. 32—Toll board messages per year in thousands, New York–Boston.

minutes. Conversations which run one half hour or an hour are not unusual, and in one case a transcontinental telephone conversation was eight hours in all.

A striking feature of the long distance service is the more rapid growth of very long haul business than business of moderate length. Figures 32, 33 and 34, for example, show respectively the growth in messages for the last five years between New York and Boston, 370 kilometers, New York and Chicago, 1,380 kilometers, and the transcontinental business between New York and Chicago at one end and

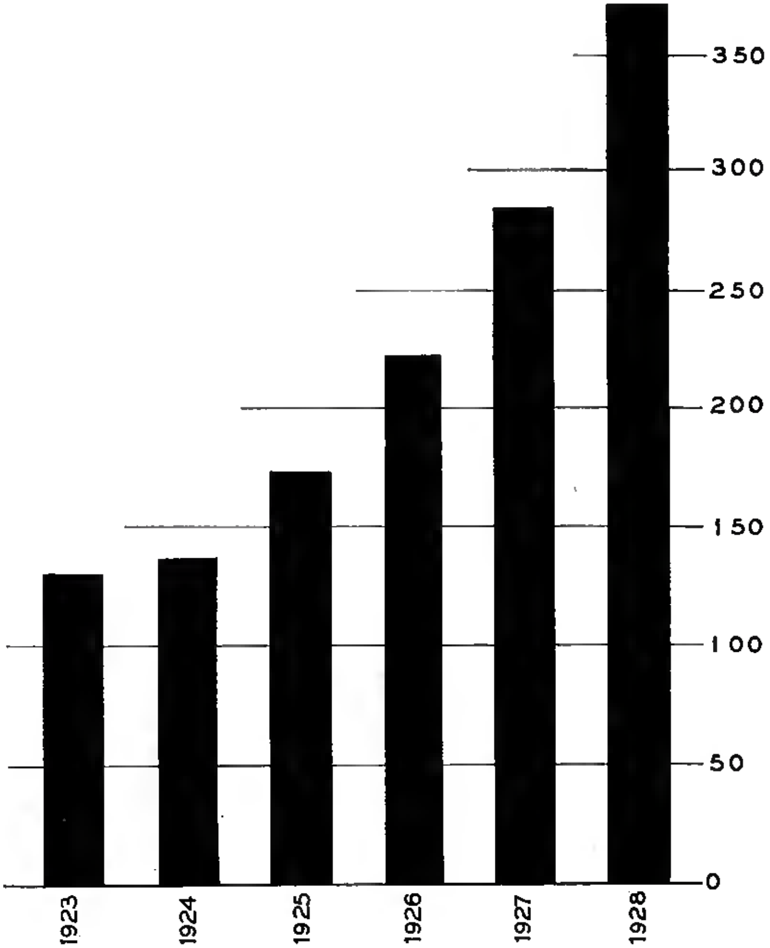


Fig. 33—Toll board messages per year in thousands, New York–Chicago.

Los Angeles and San Francisco at the other end, an average of 4,700 kilometers. It will be noted that while the toll business as a whole has increased as noted above, 67 per cent in this period, the New York-Boston business has increased 62 per cent, the New York-Chicago business 194 per cent and the transcontinental business 430 per cent.

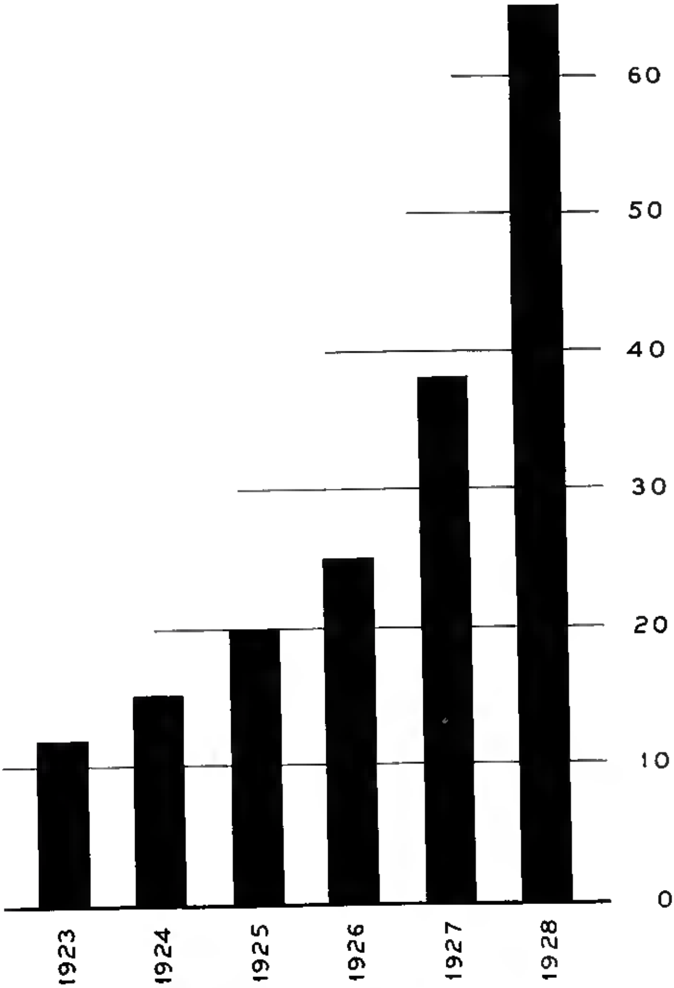


Fig. 34—Toll board messages per year in thousands, New York and Chicago to Los Angeles and San Francisco.

In the attached statistical summary is given the basis used for determining long distance toll rates, including the practices in the

United States in the offering of reduced rates in evening and night hours. The effect of these reduced rates is in some cases temporarily to slow down service at the hours when the reduced rates first go into effect, because of the large demand for long distant business at those hours.

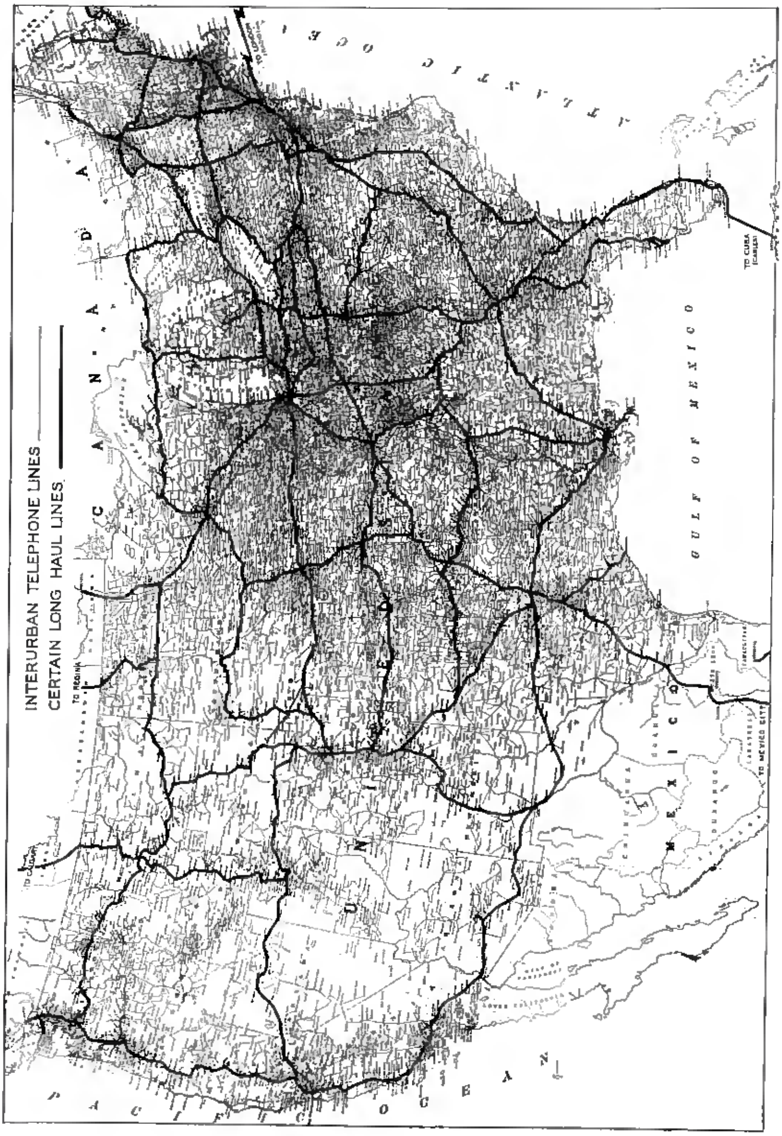


Fig. 35—Toll lines of the Bell Telephone System.

Telephone Toll Lines

To handle the toll business of the United States has required completing a network of toll telephone lines completely covering the country. This network is shown in a general way in Fig. 35. It consists at the present time of about 14 million kilometers of wire on about 300,000 kilometers of toll route. The toll circuits are partly open wire, supported on insulators and are partly in cable. Both the open wire and cable circuits are, in general, phantomed, giving three independent circuits for each two pairs of wires, and in addition on the open wire is superposed a considerable amount of carrier current telephone circuits. The distribution at the present time between

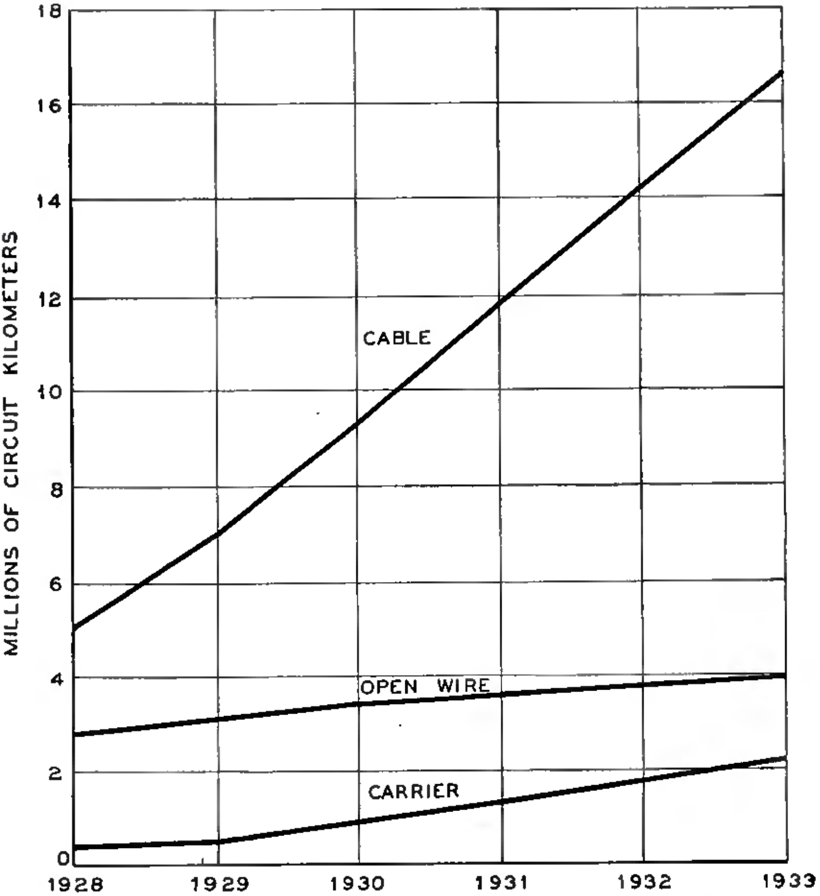


Fig. 36—Estimated toll circuit kilometers in plant Bell Operating Companies.

cable, open wire and carrier and the expected increase of each during the next four years is shown in Fig. 36.

Open Wire and Carrier Circuits

The standard construction for open wire telephone circuits in the United States is indicated by the diagram of Fig. 37 and a typical pole line built in accordance with this construction is shown in Fig. 38.

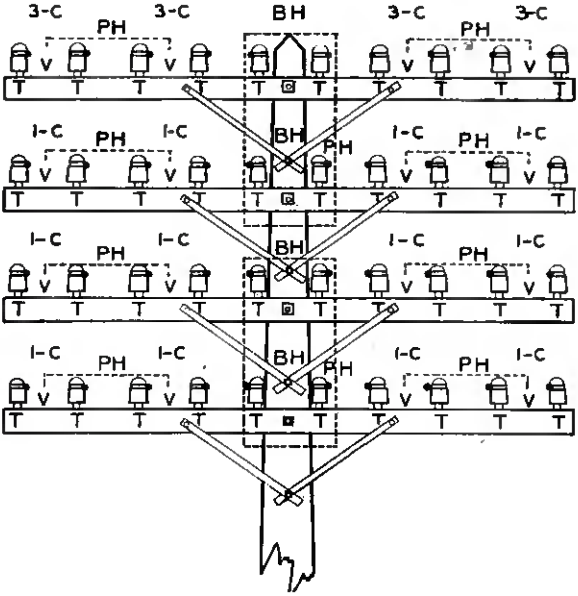


Fig. 37—Pole Line Configuration, phantomed construction, 12 inch spacing between wires of non-pole pairs.

Symbol	Facility	Total Circuits	
V	Voice frequency-physical	20	
PH	Voice frequency-phantom	10	
3-C	Carrier system furnishing 3 telephone circuits	12	
1-C	Carrier system furnishing 1 telephone circuit	12	
T	D-C telegraph	40	
BH	Carrier telegraph (10 channel)	40	
	Total telephone		54
	Total telegraph		80

The wires are of copper and the sizes and weights are shown in the following table.

Diameter—mm.	Weights—kg. per km.
2.6	47
3.2	74
4.2	118

Bronze and aluminum are not, in general, used in the United States for telephone lines, being not as economical or as generally satisfactory as copper, taking into account transmission efficiency and construction conditions.

The wires are placed on 10 pin cross-arms and supported, in general, by double-petticoated glass insulators. The grouping of wires to form phantoms is indicated in Fig. 37. This arrangement has been found desirable for conditions in the United States and transposition systems have been designed by which are obtained satisfactory operation of the phantoms and side circuits, the mutual induction between the various circuits being sufficiently neutralized to prevent mutual interference.

Fig. 38—Open wire pole line construction. Four 10-pin cross arms.

On the longer circuits telephone repeaters are installed at an average distance of about 175 to 300 kilometers, providing in that way for adequate transmission efficiency.

The number of circuits which it is practicable to provide by means of open wire lines has during the past decade been very greatly increased by the extensive use of carrier telephone for superposing on the telephone circuits additional channels of communication carried by currents above the voice range of frequencies. These systems now form a network covering the entire country and in some areas a large proportion of the circuit growth on open wire lines is taken care of by carrier systems. The systems range in length from a minimum of 75 kilometers to a maximum of 3,800 kilometers.

Two types of carrier telephone systems are standard for use in the United States. One of these, designed for the longer hauls, provides on one pair of wires three telephone circuits in addition to the voice frequency circuits. These three carrier circuits are provided by the modulation of frequencies between about 6,000 and 28,000 cycles, different frequencies being used for transmission in the two directions. The different conversations are amplified together in a common repeater at intermediate points and at the terminals separated by electrical filters providing for each circuit a band of approximately 3,000 cycles. The frequency allocation for two varieties of

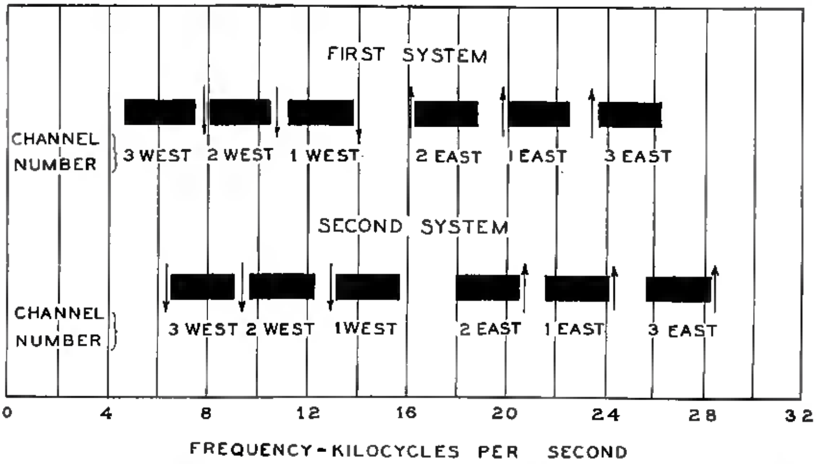


Fig. 39—Frequency allocation of two long haul carrier systems. The blocks indicate the range of the transmitted side band. The arrows are located at the carrier frequencies and indicate the direction of transmission.

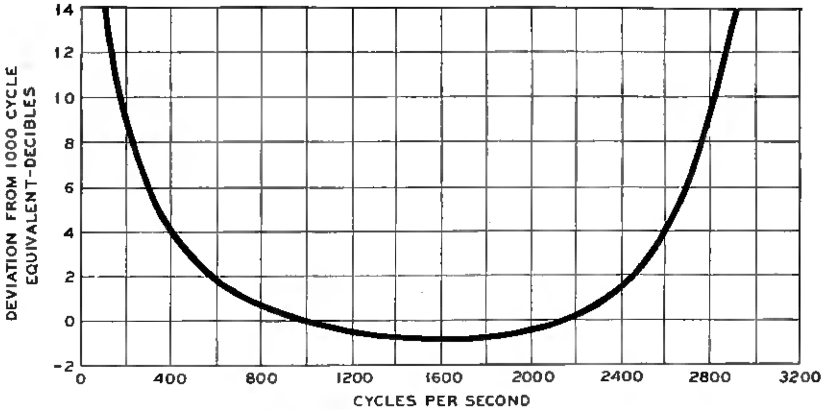


Fig. 40—Average overall transmission-frequency characteristic of Long Haul Carrier telephone system.

the long haul system in common use are shown in Fig. 39 and typical transmission characteristics for a carrier channel are shown in Fig. 40.

The long haul carrier systems give very satisfactory service and form a part of some of the longest circuits in the country. For example, the direct circuits between New York and Los Angeles, California, 5,100 kilometers in length, are made up of cable circuits from New York to Pittsburgh connected permanently to a Pittsburgh-St. Louis carrier system, and a St. Louis-Los Angeles carrier system. These two carrier systems connected together total 4,550 kilometers in length with 13 intermediate repeaters. Similarly the New York-San Francisco circuits are in cable from New York to Chicago and there permanently connected to the Chicago-Sacramento carrier system 3,800 kilometers long with 10 intermediate repeaters.

The short haul carrier system is similar in its general characteristics but is simplified and provides a single carrier circuit for each pair of wires. In the case of both systems, single side-band carrier suppression circuits are used.

In Fig. 37 showing the standard arrangement of open wires on pole lines are indicated the carrier telephone channels and also the carrier telegraph channels which can be superposed on these circuits without mutual interference, after the installation of suitable transpositions which have been designed to neutralize the mutual induction between the circuits. It is noted that with this arrangement it is possible to obtain from 40 wires 54 telephone circuits. Also 80 telegraph circuits are obtained, used for special contract service as described later in this paper.

On a number of the open wire toll routes carrying very long circuits, it has become important to provide arrangements for using a larger number of long haul carrier telephone systems, thus obtaining a larger number of circuits. Whereas a number of arrangements using the standard spacing of wires have been tried out, it is found extremely difficult to continue the use of phantoms and to so transpose the wires as to provide adequate freedom of interference between the higher frequency carrier channels if these are used on all pairs. The difficulty of doing this is evident in considering that in order to avoid overhearing it is necessary that the power transfer between different circuits should not exceed one part in a million even though they are parallel to each other for long distances.

In order to make possible the maximum use of long haul carrier systems where desired, trials have been made with the arrangement of wires shown in Fig. 41 and these trials have shown very satisfactory results. With this arrangement the spacing of the two wires of each pair except the pole pairs is reduced to 23 centimeters and the phantoms on these wires are abandoned. Type "C" systems can then be used on all of the pairs with this spacing. The result as indicated on the diagram is that a 40 wire toll line provides 70 telephone and 80 telegraph circuits.

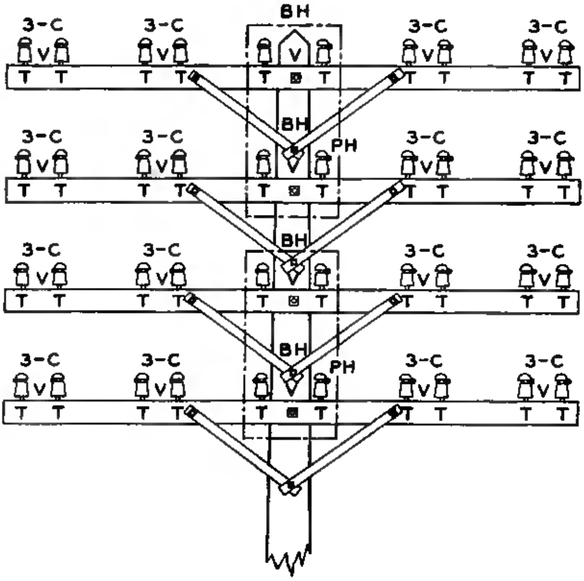


Fig. 41—Pole Line Configuration, non-phantomed construction, 8 inch spacing between wires of non-pole pairs.

Symbol	Facility	Total Circuits	
V	Voice frequency-physical	20	
PH	Voice frequency-phantom	2	
3-C	Carrier telephone	48	
T	D-C telegraph	40	
BH	Carrier telegraph (10 channels)	40	
	Total telephone		70
	Total telegraph		80

Toll Cables

In spite of the great extension in the use of open wire circuits brought about through the application of carrier systems, as indicated above, it would be extremely difficult with the present rapid growth in toll business to provide by open wire toll lines the large numbers of telephone toll circuits now required on many routes. It is very fortunate that the development of means for providing satisfactory long distance circuits through telephone cables has matured in time to enable this method of construction to be widely used to meet the present demands. Also, the toll cables provide practical immunity from the effects of storms, including the sleet storms, which are a hazard to open wire construction in nearly all parts of the United States.

The first long distance toll cables in the United States were placed in service in 1906 between New York and Philadelphia and between Chicago and Milwaukee. These cables were both placed underground in multiple duct and are each about 150 kilometers long. The next step in the extension of toll cables was the completion in 1914 of an underground toll cable route between Boston, New York, Philadelphia and Washington, a distance of 730 kilometers. Cable running west from New York was completed to Chicago, a distance of 1,380 kilometers, in 1925 and to St. Louis, a distance of 2,150 kilometers, in 1926. This permitted placing in service circuits entirely in cable between New York and St. Louis.

The present major toll cable routes together with the extensions which it is expected to complete during the next five years are indicated in Fig. 42. It is to be seen that in accordance with these plans toll cable will, within five years, extend entirely across the continent and up and down the length of both Atlantic and Pacific Coasts, will extend north into Canada and south almost to Mexico. In the northeastern portion of the country where the development is the heaviest, there is already a multiplicity of toll cable routes and on some of these routes the rate of growth is high enough to require additional cables at successive intervals of one or two years. The amount of toll cable added to the network this year will be about 8,000 kilometers and this amount is expected to be increased materially in the following years.

In the early toll cables before the extensive development of telephone repeaters, it was necessary, in order to provide satisfactory transmission, to use relatively large conductors and conductors up to a maximum size of 2.6 mm. diameter (No. 10 B and S gauge), were provided in the Boston-Washington cable.

With the perfection of telephone repeaters for use with toll cable circuits, the transmission limitations on the extension of toll cable were removed and the economy of such circuits greatly increased by making it possible to use small conductors. The longest toll cable circuits at the present time are carried over conductors of 0.9 mm. diameter (19 B and S gauge). For the shorter circuits each path is used as a two-way circuit, while the longer circuits use separate paths for transmission in opposite directions. In order to improve the transmission characteristics, the circuits are provided with loading coils

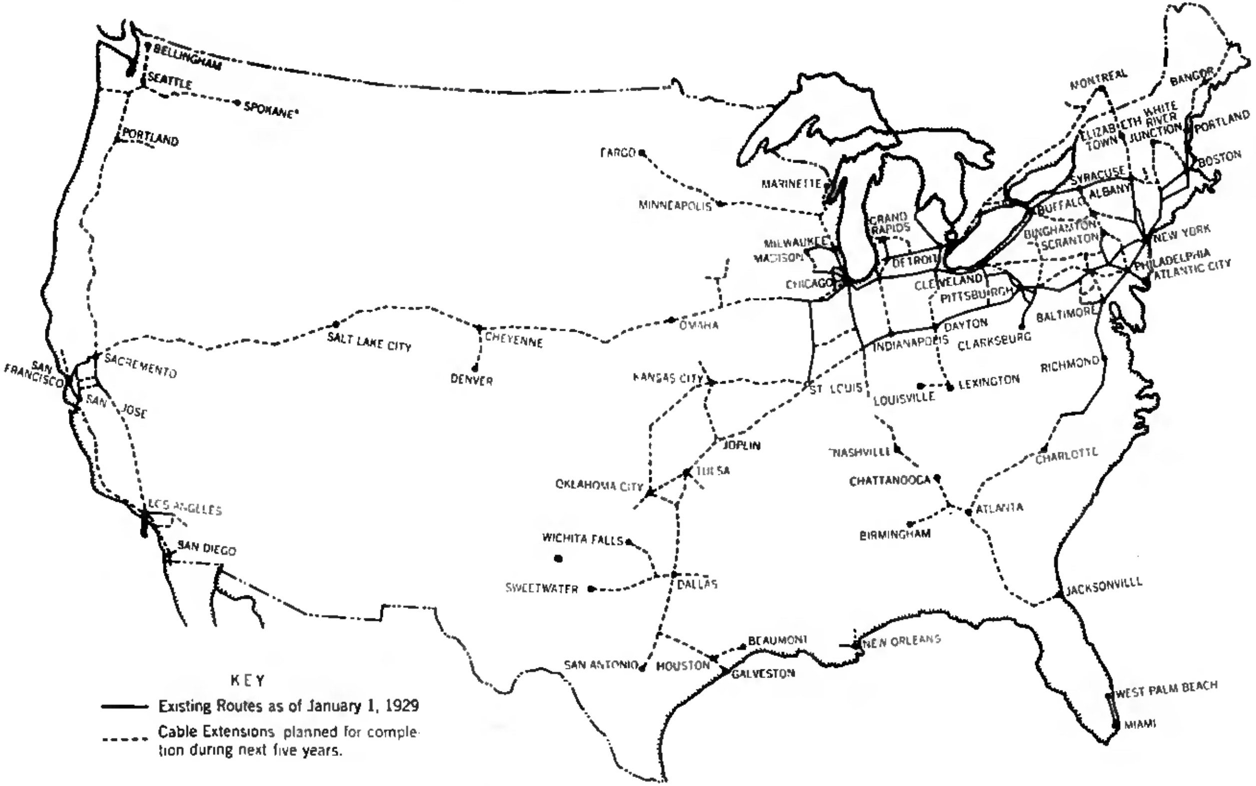


Fig. 42—Main toll cable routes of the United States.

at intervals of 1,830 meters and at an average interval of about 70 kilometers are provided with telephone repeaters which renew the power of the attenuated voice currents. A single standard full size cable 6.7 cm. in diameter when so equipped is capable of providing between 250 and 300 long distance telephone circuits.

The toll cable system includes various types of construction. For the routes having the most rapid growth, multiple duct subway is used. At the present time with the development of very heavy toll demands in many parts of the country, this type of construction is being extended very rapidly on a number of important routes. Multiple tile duct with small splicing manholes located at intervals of 229 meters and large manholes for loading coil pots at intervals of 1,830 meters are generally used.

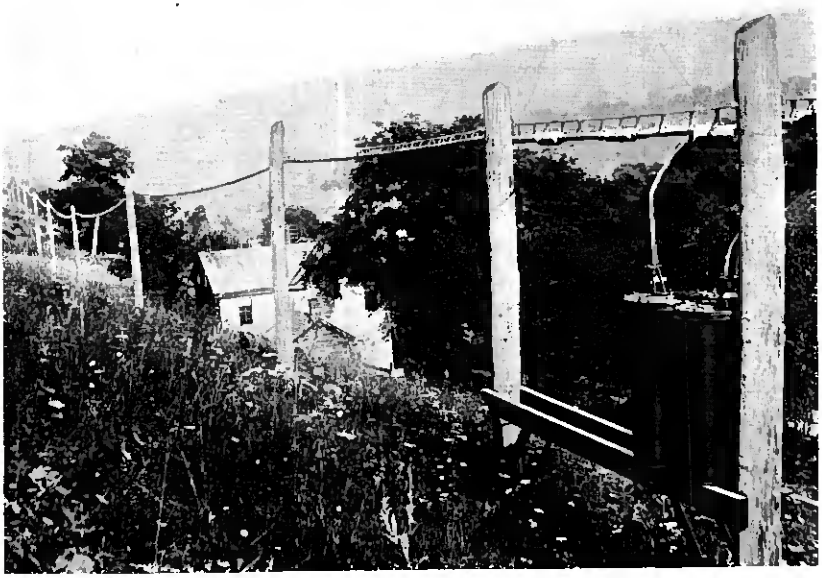


Fig. 43—Typical aerial toll cable construction showing loading point.

For routes on which the growth is relatively light, for example, 40 or 50 circuits a year and where underground construction is desirable, two other types of construction have been used to a limited extent. In one type the cable is placed in a single duct of fibre and in the other type of construction cable covered with a double layer of steel tape is placed directly in the earth. With both of these types of construction, manholes are built only at loading points.

In many places the character of the country is such that underground construction would be very expensive. In such cases, and in other cases where it seems desirable, aerial toll cable construction has been used extensively in the United States. With this type of construction the cable is suspended from a steel messenger wire supported on poles. Figure 43 shows typical aerial cable construction, including a loading point, the pots of loading coils being supported on an angle iron pole fixture.

Long circuits in toll cables have some extremely interesting electrical characteristics. Figure 44 shows the net transmission characteristic over a range of frequencies of a New York-Chicago toll cable circuit 1,380 kilometers in length. It will be seen that the voice frequencies are transmitted with nearly the same net efficiency over a sufficiently wide band to give a high grade of telephone transmission. The net characteristic indicated, however, is obtained by almost wholly neutralizing with telephone repeaters the very large transmission loss in the circuit. The New York-Chicago circuit, for example, would have an attenuation loss at 1,000 cycles of about 470 db, which means that without amplification the ratio of output power at one end to input power at the other end of the circuit would be 10^{-47}. The combined gain of the 19 telephone repeaters in the circuit is about 461 db, giving about 9 db net equivalent. Under these conditions, it is evident that a careful regulation of the circuit is essential. For example, variations in the temperature of a circuit in the course of a day could make as much as 30 db or 1,000 fold difference in the

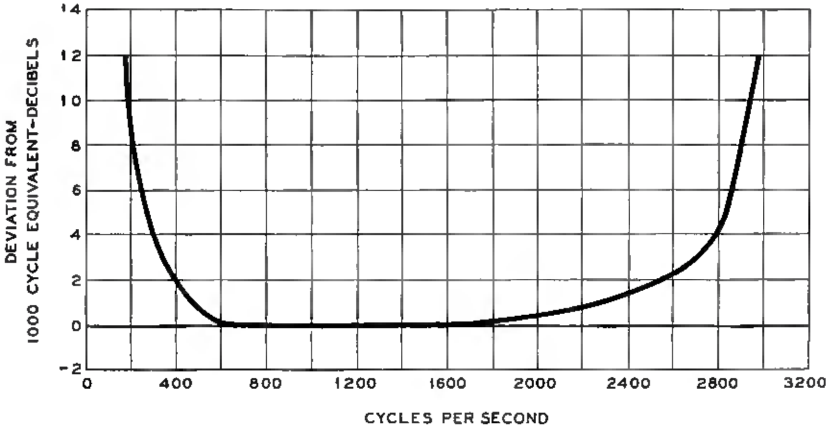


Fig. 44.

electric power received at the end of the circuit. To prevent such variations affecting the net equivalent the long circuits are all provided with automatic regulators which adjust the gains of the telephone repeaters to compensate for the effect of temperature variations on the equivalent of the circuit.

The effects of transmission delay are also very interesting and important. Voice waves travel considerably more slowly over cable circuits than they do over open wire circuits. For example, the velocity is about 30,000 kilometers per second for "long distance" type cable circuits as compared to nearly 300,000 kilometers per second for non-loaded open wire circuits.

One important result of delaying the speech waves is the "echo" effect. The transmitted currents are in part reflected at the distant terminal due to variations in the impedance of the receiving circuit. If the reflected currents transmitted back to the other end are delayed enough they may be heard by the talker as echoes of his voice. They may be again reflected at the sending end of the circuit and returned to the listener as an echo following the directly transmitted speech. The effects of these echoes are largely eliminated by devices known as "echo suppressors" by means of which the transmission of voice waves in one direction over the circuit causes interruption of the path over which the echo currents are transmitted in the opposite direction. However, the effectiveness of echo suppressors is limited by the necessity that they shall not be operated by noise currents of extraneous origin as this would interrupt conversations. The echoes, therefore, are an important factor to be taken into account in determining the type of toll cable circuit to be provided to meet the transmission limitations imposed on the long distance circuits.

In cable circuits introducing considerable transmission delay, the fact that the delay is not exactly the same for waves of different frequencies is also important, tending to give rise to what have been sometimes referred to as "transient" effects. In loaded cable circuits the waves of higher frequency are delayed more than those of lower frequency because of the fact that the loading is applied in lumps. The coils and condensers in the repeaters and auxiliary apparatus, on the other hand, tend to delay the waves of lower frequency. The result is that the waves of intermediate frequency arrive first, followed by the waves of higher and lower frequency. Devices known as "phase compensators" can be used to reduce the effects, particularly those caused by the line. To improve the situation at the low end of the frequency scale special attention has been given to the design of the repeaters and auxiliary apparatus.

Still another effect of the transmission delay is to somewhat slow up and perhaps otherwise interfere with conversations due to the delay which is added to the ordinary time elapsing between question and answer. For example, if a cable circuit is 5,000 kilometers long and the voice waves travel 30,000 kilometers per second, the time required for the waves to travel from one end of the circuit to the other is $\frac{1}{6}$ second and to make a complete round trip, $\frac{1}{3}$ second. This $\frac{1}{3}$ second delay is evidently added to the ordinary time which elapses between question and answer. In the United States cable connections somewhat longer than 5,000 kilometers will be used in

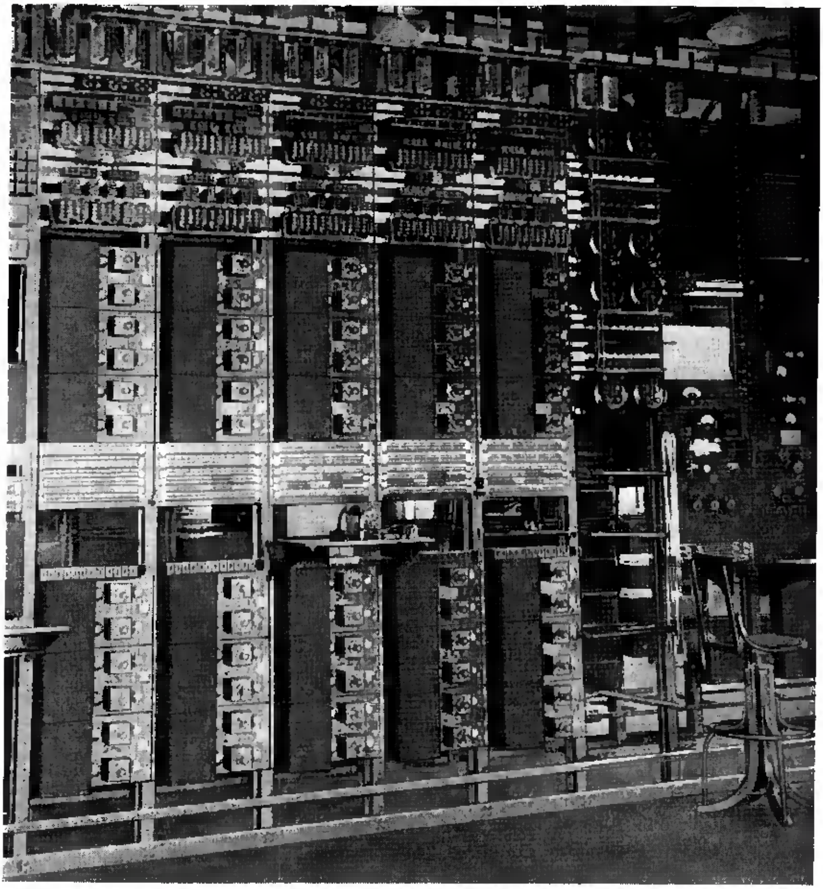


Fig. 45—Thirty 4-wire repeaters and associated testing equipment. The repeaters are arranged in groups of 3 with a minimum of cabling, each group being associated with a phantom circuit and its 2 side circuits.

the future, while for international connections, of course, very much longer distances than this will be involved. In the United States considerable study is, therefore, being given to the effects of transmission delay and to methods of avoiding difficulties on the very long connections including the development of cable circuits of higher speed.

The toll cable circuits today include two principal types, one, discussed above, for the longer distances having a transmission speed of about 30,000 kilometers a second, and the other for the shorter distances, transmitting a narrower band of frequencies and having about one-half the transmission velocity. In view of the superior transmission characteristics of the long distance type circuits it is the present practice in the design of new toll cable circuits in the United States to limit the use of the short distance type facilities to circuits about 160 kilometers in length if they are to be used for switched business, and about 280 kilometers in length if used only for terminal business.

Toll Circuit Equipment

The apparatus required for the operation of toll circuits has been developed in the form of panels mounted on standard bays of angle iron, thus bringing about a great reduction in the space required compared with earlier forms of mounting. Figure 45 shows a bank of 30 four-wire repeaters arranged in groups of three, each group being associated with a phantom and its two side circuits. Figure 46 shows the panels containing complete terminal equipment for two type "C" carrier telephone systems (six circuits) with associated testing apparatus.

The equipment is housed in fire-proof buildings. Figure 47 shows a typical telephone repeater station, this one being located at Princeton on the cable route between New York and Philadelphia. This building now contains 1,100 repeaters. Some of the telephone repeater stations now being built are designed for ultimate capacity with extensions of 10,000 repeaters.

An interesting feature of the long telephone circuits is the use of 1,000-cycle current for signaling rather than the lower frequencies which have been general in the past. This higher frequency has the advantage of being efficiently transmitted by the telephone circuit without change in the amplifying apparatus and hence does not require intermediate ringing apparatus. At the terminals it is rectified and caused to operate relays which actuate the desired signal.

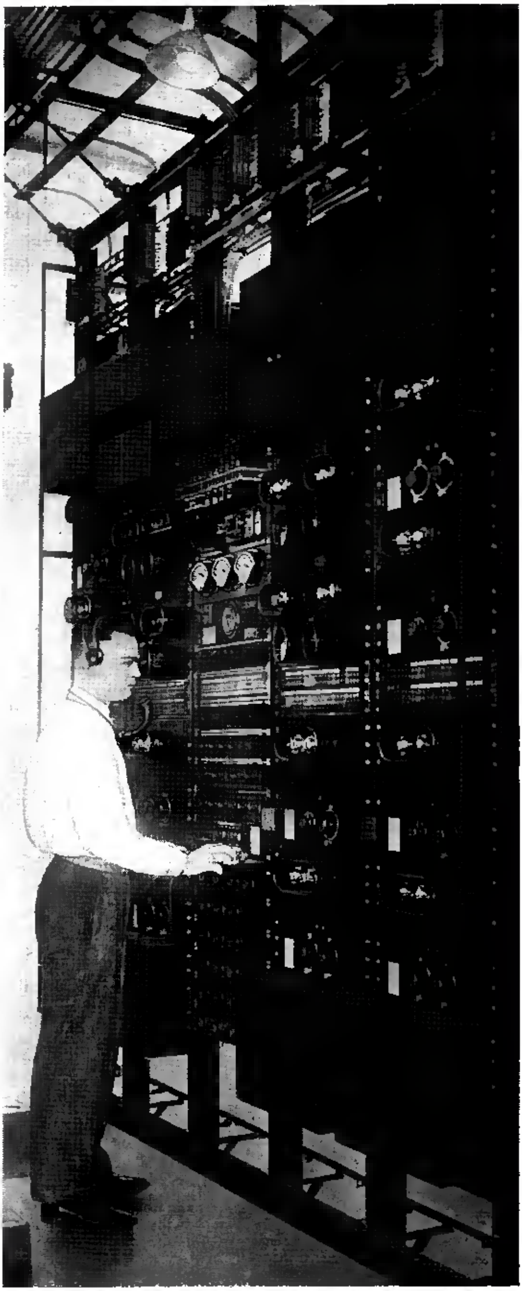


Fig. 46—Complete terminal repeater apparatus for two long haul carrier telephone systems (6 circuits) with associated testing equipment.

Fig. 47—Telephone repeater building at Princeton, New Jersey on New York–Philadelphia cable route. Building now houses 1100 repeaters. Ultimate capacity 2200 repeaters.

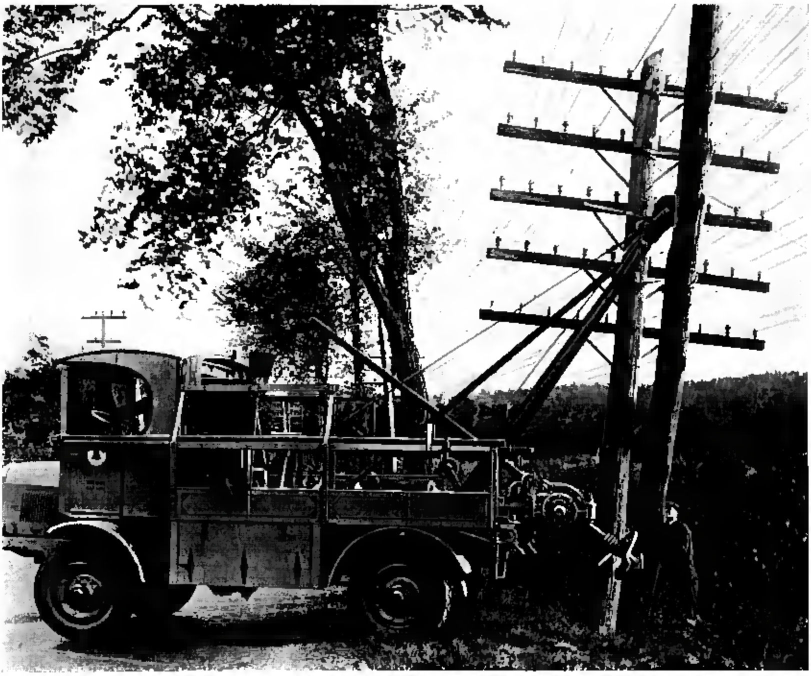


Fig. 48—Earth boring machine and derrick. Will bore 60 centimeter hole 2 meters deep in loam or clay soil in about one minute and in stone or frozen soil in 5 or 10 minutes. Derrick operated by power driven winch for setting poles. Truck provided with four wheel drive.

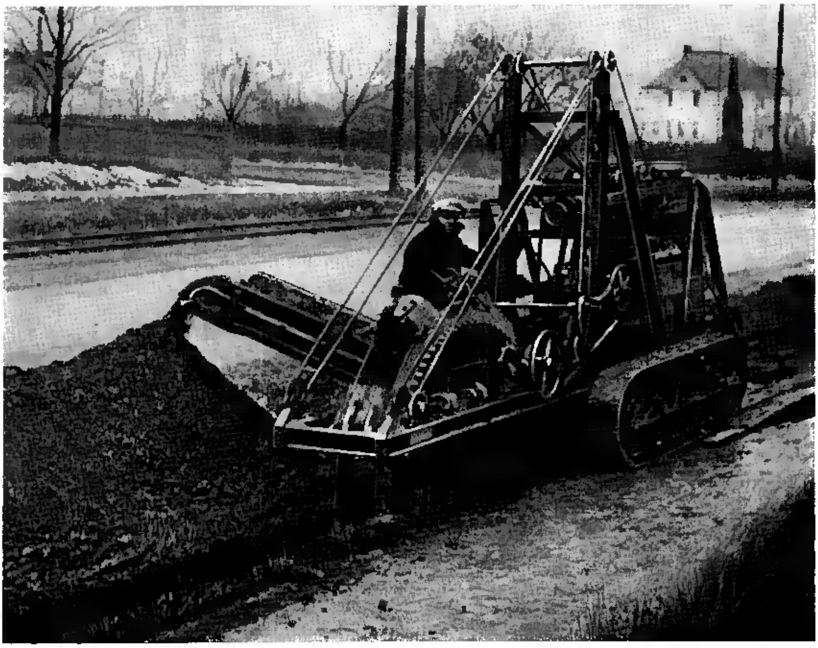


Fig. 49—Trenching machine. Digs trench 1.7 meters deep and 55 centimeters wide, at speeds varying between 0.2 and 1.2 meters per minute, and is carried from job to job upon a trailer drawn by 2½ ton truck.

Fig. 50—Automobile truck equipped with tracks for hauling cable on private right of way. With tracks, speed about 16 kilometers per hour. Can carry 4500 kilogram reel up 40% grade. Tracks can be removed using special equipment provided for that purpose; without tracks speed 27 kilometers per hour.

Toll Line Construction

The construction of toll lines under a wide variety of conditions has required the solution of many interesting problems. The relatively high cost of labor in the United States contributes to the extensive use of labor saving machinery, a large amount of which has been developed to meet the particular conditions of telephone construction. Figures 48, 49 and 50 illustrate some of the more interesting types of labor saving machinery used extensively for both open wire and toll cable construction.

Numerous special construction problems are, of course, met in specific situations. One of the interesting river crossings is illustrated in Fig. 51.

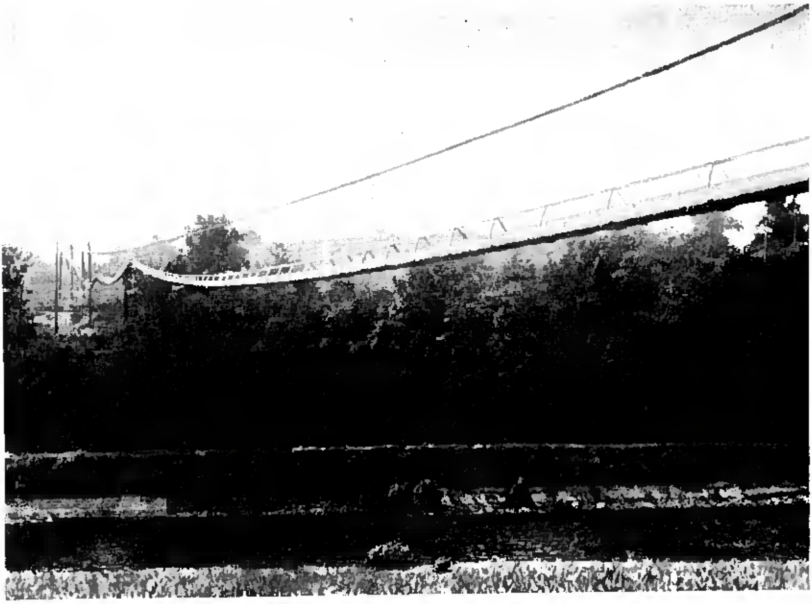


Fig. 51—Special aerial cable construction across a river. Cable and messenger secured to a catenary suspension wire. 2-spans each about 180 meters long.

Switching of Toll Circuits

As far as is economically practicable the toll business is handled by direct circuits without intermediate switching. At the present time this includes 80 per cent of the toll messages. Of the remaining 20 per cent, 17 per cent have one intermediate switch and 3 per cent more than one intermediate switch.

It is the purpose of the Bell Telephone System to design the toll

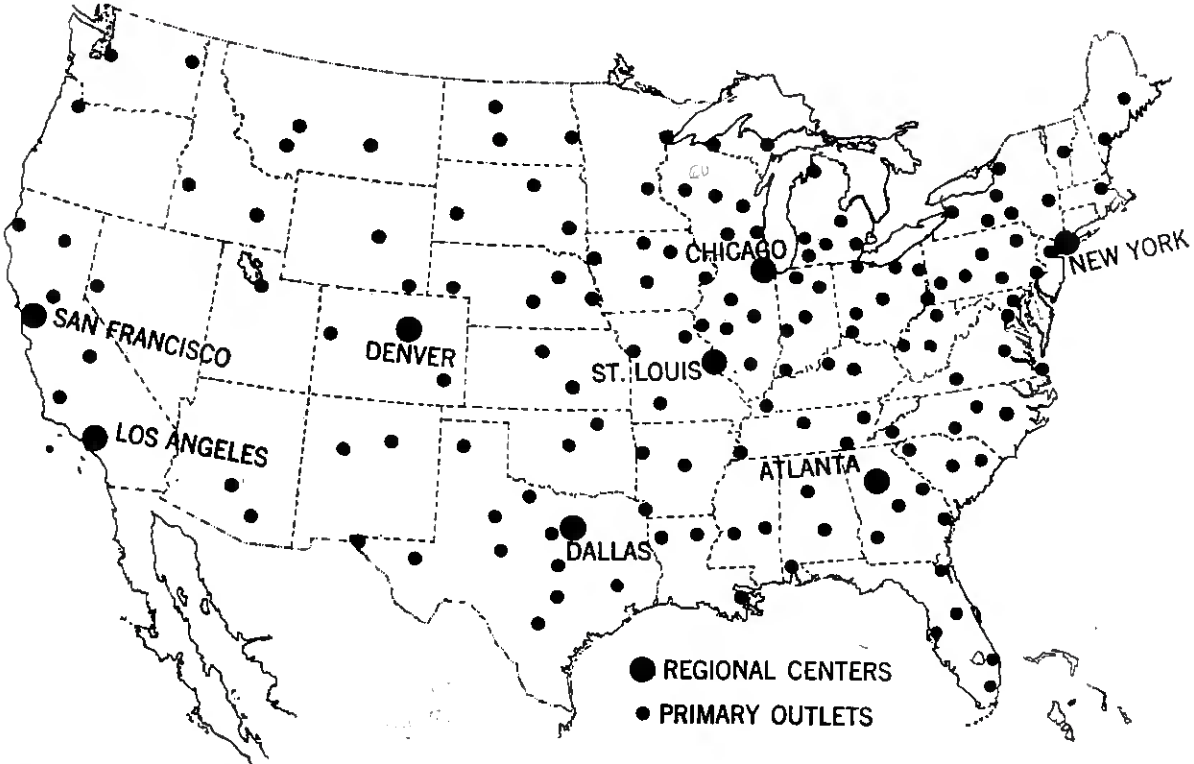


Fig. 52—General toll switching plan. Map showing location of important switching points in Bell Operating Companies.

telephone system in the United States to give satisfactory service between any two points in the country. In order to accomplish this it is necessary to make arrangements for a minimum number of switches between any two points. Also the toll circuits which will be used as parts of the built-up connections must be designed for a very high standard of transmission so that the overall efficiency of the built-up connection will be satisfactory.

Arrangements have recently been worked out in the United States for meeting requirements of switched traffic more satisfactorily than has heretofore been possible. These arrangements may be briefly described as follows:

At different points in the country there have been selected a group of eight very important switching points shown in Fig. 52. These eight regional centers will all be interconnected by high grade groups of circuits directly, that is, without intermediate switch. Throughout the country there are selected about 147 important switching points known as primary outlets also indicated in Fig. 52, each of which is directly connected to at least one of the regional centers. Each of the remaining 2,576 toll offices in the country will be connected to at least one of these important switching points. Furthermore, within limited areas, such for example as a State, all important switching points will be directly interconnected. Within such an area, therefore, any two toll offices can be connected together with not more than two intermediate switches. Also, every toll office can be connected to a regional center with not more than one switch and through that center can reach any other toll office in any part of the country with a minimum number of switches.

To insure adequate transmission on the switched connections, each of the important switching points will be provided with means for automatically inserting gain in the connection when two toll circuits are switched together so that the overall connection may be operated at the highest possible efficiency. This will, in general, be done by automatic adjustment of the gain of terminal repeaters permanently installed in circuits which must, in general, because of transmission limitations be operated at a lower efficiency when used for terminal business than when used as parts of a built-up connection.

Maintenance of Toll Service

With the present network of long distance lines in the United States, it is common to have 20 or more repeaters installed on each of the longer circuits and this number will increase greatly with the further extension of toll cable. The maintenance of service over these long and complicated circuits is a very considerable problem both from the standpoint of technique and of organization. In this paper, these problems will not generally be discussed, but certain features will briefly be indicated.

The service maintenance of the circuits includes periodic tests of transmission efficiency with transmission measuring sets designed for rapid and efficient use by the plant maintenance forces. The frequency of tests varies according to the requirements of each circuit group.

To expedite the testing and adjustment of the circuits the longer cable circuits are subdivided into circuit units, these units usually being in cable about 160 to 240 kilometers in length, including the conductors and equipment involved in one section arranged for the automatic compensation of temperature variations. When trouble occurs on a long circuit, the circuit unit in which the trouble is located is immediately replaced and the location of trouble within the circuit unit then can be carried out without further interruption of service. The responsibility for establishing and maintaining each circuit group is given to a control office which is provided with private communication channels to all parts of the circuit.

An important feature in the maintenance of long toll circuits is the physical relations between the telephone circuits and circuits for the transmission or distribution of electric power. The Bell Telephone System and the power companies of the United States as represented by the National Electric Light Association are very actively cooperating in a study of the best means of so coordinating the plant of telephone and power companies as to avoid interference under the various types of conditions important in practice. By means of this work it has been possible to find in every case a satisfactory solution permitting each utility to extend and increase its service along natural lines and providing proper protection of the telephone service.

INTERNATIONAL CONNECTIONS

General

The connections between the telephone systems of the United States and the telephone systems of other countries are indicated in Fig. 53.

The territory of the United States has direct contact with only two other nations, Canada on the north and Mexico on the south. The common language and the close commercial relations between Canada and the United States have naturally resulted in a well developed arrangement of lines connecting the telephone systems of the two countries. Telephone connection between the cities of the United States and Mexico was not made until 1927, due to the unsettled political conditions which obtained for some years in Mexico.

The many close commercial, political and social relations between the peoples of Europe and America have naturally drawn the attention of telephone men for many years to the possibility of establishing telephone communication between these two continents. It was a great satisfaction, therefore, to be able to inaugurate such a service in 1927. The transatlantic telephone circuits already connect over 20,000,000 telephone stations in North America to over 7,000,000 telephone stations in Europe, thus joining together over 85 per cent of the total telephone stations of the world.

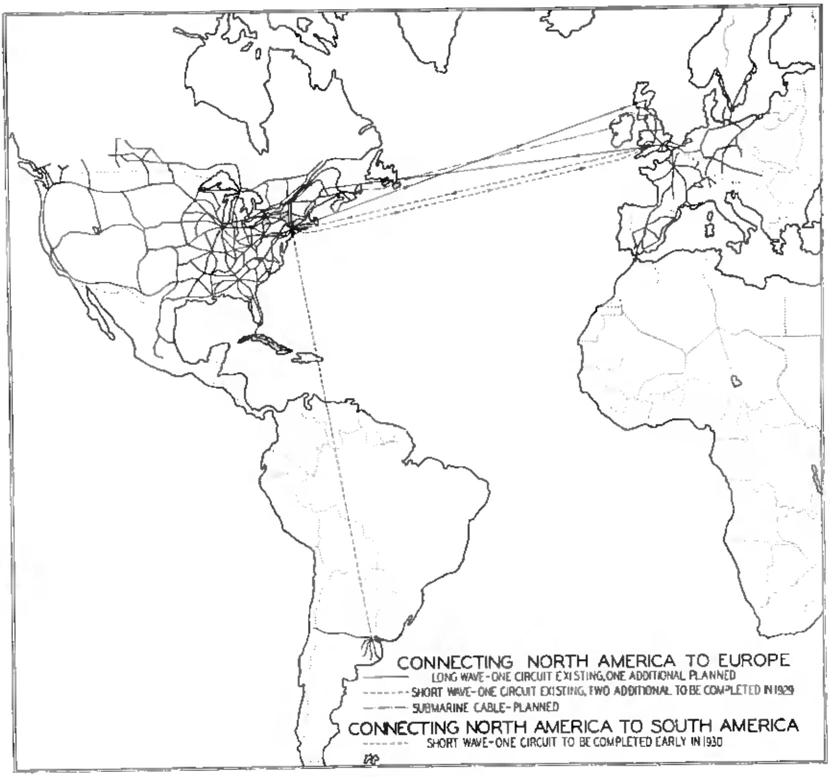


Fig. 53.

In somewhat more detail, the present status of the connections of the United States telephone system to the telephone systems of other countries is as follows:

Connections in North America

Practically all the telephone stations in Canada have communication to the telephone stations in the United States. There are approximately 100 long distance circuits extending from cities in the United States to important Canadian centers, including Halifax, St. Johns, Montreal, Toronto, Hamilton, Winnipeg, Regina, Calgary and Vancouver. The remaining cities are reached either directly or by switching through the important centers. In addition to long distance circuits there are, of course, many short distance circuits connecting points on opposite sides of the boundary which have local relations with each other. The various companies and provinces in Canada coöperate very closely with the Bell Companies in the United States in the maintenance of international service and, in general, telephone practices are very similar or identical in the two countries.

Telephone communication is extended from the United States to Mexico by means of a telephone line crossing the border near Laredo, Texas. Direct long distance circuits extend from points in the United States to Mexico City, Tampico and Monterey and through these centers to about one-half the telephone stations in Mexico. Local toll circuits cross the border at a number of places.

Telephone communication was established between the United States and Cuba in 1921 by the placing of three telephone cables between Key West and Havana. Each of these cables furnishes one telephone circuit and a maximum of four telegraph circuits. The requirements for the cables were exacting since a length of about 190 kilometers is combined with a depth of water having a maximum of 1,860 meters. Each cable consists of a central conductor magnetically loaded with a wrapping of fine iron wire and insulated with gutta percha compound. A metallic return path for the telephone currents is furnished by heavy copper tape wrapped outside of the insulation and, therefore, in contact with the surrounding water. Three of the telegraph circuits in each cable are obtained by using "carrier currents" at frequencies slightly above the voice range. The fourth is obtained by using frequencies below the voice range.

Connections to Europe

In 1915 the Bell System experiments on radio reached the point where telephone messages were transmitted by radio from the United States and were successfully received by engineers sent for the purpose to Paris and to the Hawaiian Islands. While the Great War delayed technical and commercial development, in 1923 the Bell Companies were able to carry out a successful demonstration of radiotelephone transmission from a group of telephone officials in New York to a group of people interested in communication assembled for the purpose in London. The success of these experiments led to cooperation with the British Post Office and the establishment in 1927 of telephone service between New York and London. This service has now been extended to include the greater part of the telephones of North America and Europe.

As indicated in Fig. 53 there now exist one long-wave and one short-wave telephone circuit between the two continents. A second short-wave circuit will be placed in service about June 1 of this year and a third in December. By the end of 1933 it is expected that there will be in service between New York and London a group of six circuits consisting of three short-wave radio circuits, two long-wave radio circuits, and one cable circuit. Our best information indicates that the short-wave circuits will be suitable for service at least 60 per cent of the time, the long-wave, 90 per cent, and the cable, 100 per cent.

Since the beginning of 1929, the average number of messages handled per week has been 275. For this period the average number of messages per day, omitting Saturday and Sunday, has been 44. Eighty-nine messages were handled on Christmas Day, 1928.

Certain technical features of these circuits are particularly interesting. The long-wave circuit operating at a frequency of approximately 60,000 cycles employs the "single side-band carrier suppression" method. This appears to be the only use of this method in radio, although it is widely used in "carrier" circuits over telephone wires. The energy saved by the suppression of the carrier and the increased selectivity permitted by the narrow band of frequencies which is transmitted gives this system a transmission effectiveness as great as a system of three or more times as much power using the ordinary transmission method. At both ends the receiving stations are situated as far north as can conveniently be reached and use is made of highly directive receiving. It is estimated that at the United States end these two factors represent an improvement equivalent to an increase in power of five thousand times as compared to a non-directive receiving station located at the same latitude as the transmitting station.

The short-wave transmitting and receiving stations located not far distant from New York and London employ highly directive antenna systems. The design of such antennas must take into account economic factors and possible reactions on receiving effects other than power efficiency such as fading. The improvements effected by such systems depend on wave-length and transmission conditions. Under favorable conditions the improvement effected at each end is approximately equivalent to a transmitted power increase of 100 times. The most useful wave-lengths for this service have proved to be in the vicinity of 16 meters, although wave-lengths of about 22 and 33 meters are also provided to increase the amount of time these circuits are satisfactory for service because at certain seasons and times of day they are more effective than the 16 meters wave-length.

Service over the transatlantic facilities is carried on from 6.30 in the morning to 10.00 at night in New York, corresponding to 11.30 in the morning and 3.00 A. M. London. During the winter months the long waves give nearly continuous service over this period. Under summer conditions considerable difficulty is frequently experienced in maintaining the long waves during the afternoon period in New York, corresponding to the evening period in London. At these times, however, the short waves are usually effective.

The projected transatlantic telephone cable will use new magnetic loading materials and new insulating compounds for submarine cables recently developed by the Bell Telephone Laboratories. It will have at least one intermediate repeater point at Newfoundland. A circuit of this kind, differing radically from radio circuit in its characteristics will add both to the message capacity and to the reliability of the transatlantic service.

Connections to South America

Figure 53 indicates a short-wave radiotelephone circuit from New York to South America which, it is expected, will be in service early in 1930. The South American transmitting and receiving stations, which will be in the vicinity of Buenos Aires, will be owned and operated by the companies who operate the local telephone service in Buenos Aires and the wire lines extending to other points in South America.

SPECIAL SERVICES

Telegraph Circuits

While the Bell System handles practically no commercial telegraph message business, it plays an important part in meeting the communication needs of the United States by furnishing a large mileage of telegraph circuits for the private use of individuals and institutions, and for the use of governmental departments. Over two million kilometers of such circuits are now in use. One-third of this amount is used by newspapers and press associations. The greater part of

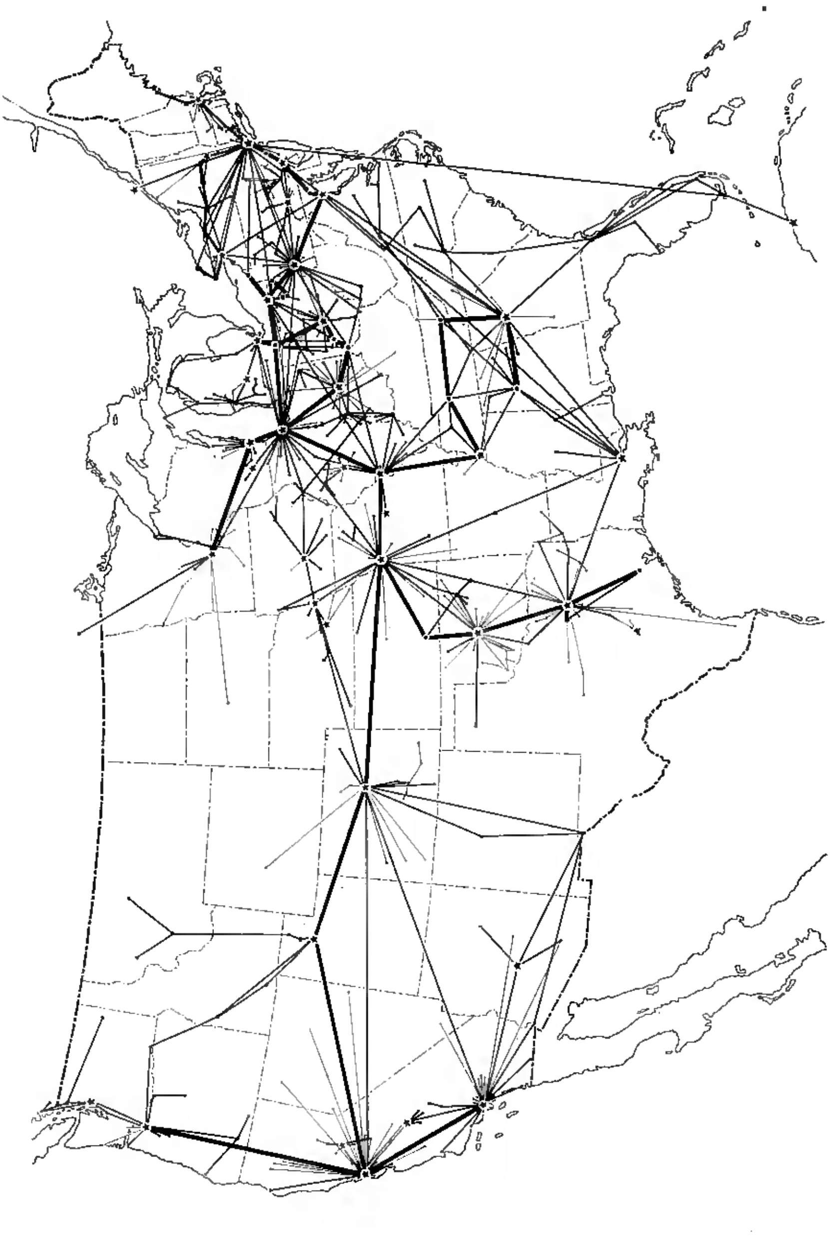


Fig. 54—Special contract telegraph service furnished a press association. 53 circuits totaling 124,000 kilometers, are used exclusively by the press association for distributing news to its customers.

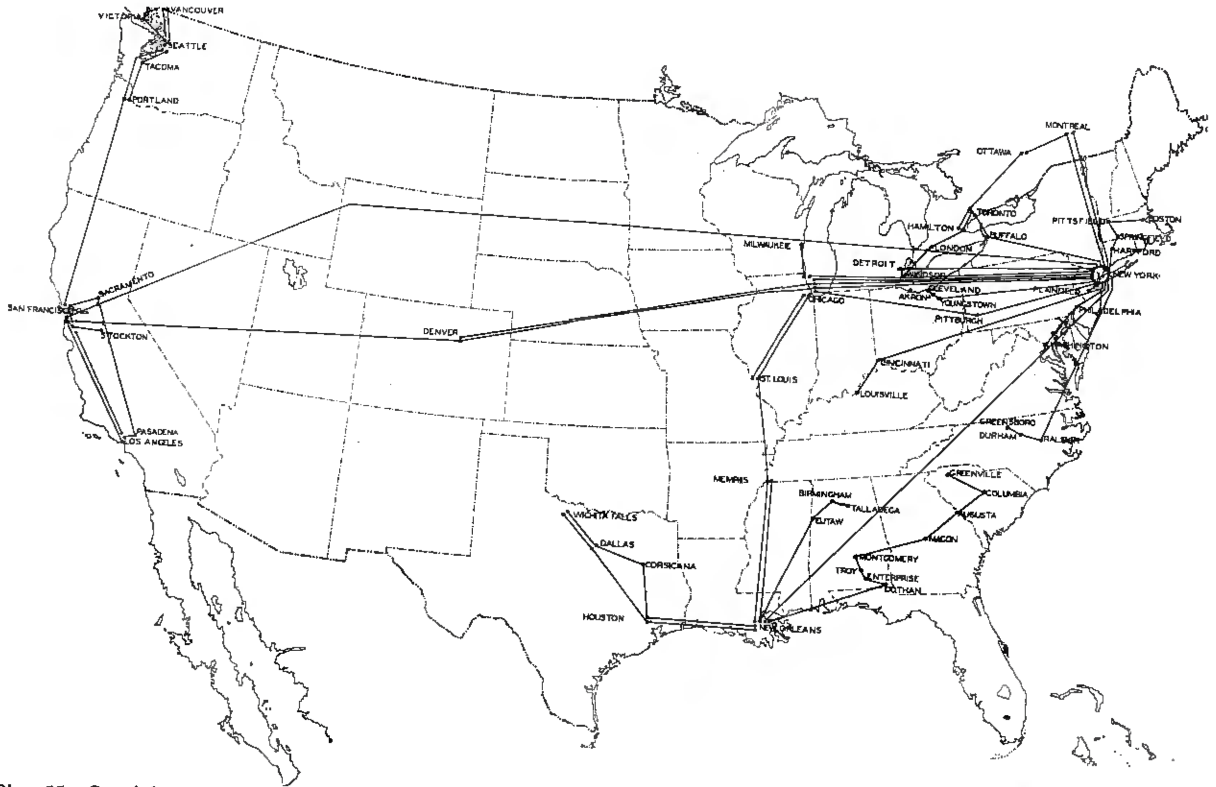


Fig. 55—Special contract telegraph circuits furnished to a brokerage company. 30 circuits totaling 38,000 kilometers with 95 stations.

the remainder is used by commercial, financial and other organizations. Between New York and Chicago, a distance of approximately 1,400 kilometers, there are slightly over 300 such circuits now in operation.

Figure 54 shows the system of telegraph circuits furnished by the Bell Companies to one of the press associations. An indication of the importance of private communication systems to commercial and financial institutions is given in Fig. 55 which shows the telegraph circuits furnished by the Bell Companies to a single brokerage company.

The greater part of such telegraph circuits have in the past been operated by hand-speed Morse telegraph. At the present time, however, nearly a third of the mileage is operated with telegraph printers and this method of operation is rapidly increasing. Two speeds of service employing printers are offered, one operating at 40 words per minute and the other operating at 60 words per minute. At the present time, in view of the use to which this service is put, no demand has arisen for multiplex operation, but this method of operation is possible and will be used if it should become desirable.

The telegraph circuits were originally all obtained as a by-product of the telephone business by compositing or otherwise superposing them on the telephone wires, using direct current for the telegraph circuits. At the present time approximately two-thirds are obtained in this way. The remaining third are obtained by "carrier current" methods. The carrier current system of open-wire lines uses frequencies above the voice range and provides ten duplex telegraph circuits on each pair of wires. The carrier current system used on cable circuits employs frequencies within the voice range, the currents being transmitted over an ordinary telephone four-wire cable circuit. This system gives twelve duplex telegraph circuits on each such circuit.

Telephone Circuits Provided for Private Use

In addition to the usual telephone message business, the Bell Companies furnish telephone circuits for the private use of individuals and organizations.

So-called "special contract" telephone circuits are set up between particular parties for their private use at definite times specified in the contract. Approximately 2,000,000 circuit km. hours of such facilities are now in use during each complete business day. This is the sum of the figures obtained by multiplying the length of each such special circuit by the number of hours per day it is continued in use.

About three quarters of this total is accounted for by circuits where the contract calls for 12 hours operation per day, nearly all the remainder is accounted for by circuits which remain in service 24 hours per day. A remaining small fraction is made up of shorter period contracts which are permitted to be as short as 30 minutes per night one night per week, or 10 minutes per day five days a week.

As an illustration of the extent of use of this service, there are at present 158 full-time special contract circuits between New York and Philadelphia and 89 of such circuits between New York and Boston.

Foreign Exchange Service

Closely related to the above is the furnishing of what is called foreign exchange service. This consists of an arrangement whereby a customer in one exchange area is provided with a circuit for his exclusive use to another exchange area, this circuit being associated with a telephone number in a distant exchange so that other telephone stations in that exchange can be connected to the special line without toll charge. By this means, a business office in Boston, for example, can be given a New York telephone number, all New York calls for that number being treated as local calls but being actually completed over the special line to Boston.

This type of service has a considerable popularity, there being over 1,000 such lines in service at the present time. Most of them are for relatively short distances, but some are for material distances, the longest being between Cleveland and New York, a distance of about 900 kilometers.

Telephone Networks for Program Transmission to Radio Broadcasting Stations

Radio broadcasting has resulted in the development of networks of telephone circuits for transmitting programs from studios or other places at which they are picked up to the radio station or system of stations from which they are broadcast. By such telephone wire systems the ceremonies of the Presidential Inauguration on March 4, 1929, were simultaneously transmitted to 118 radio stations located all over the United States. A statement regarding these interesting telephone networks, the requirements which they must meet and their importance in program broadcasting in the United States is given in a separate paper presented to this Congress (see paper on Wire Systems for National Broadcasting by A. B. Clark).

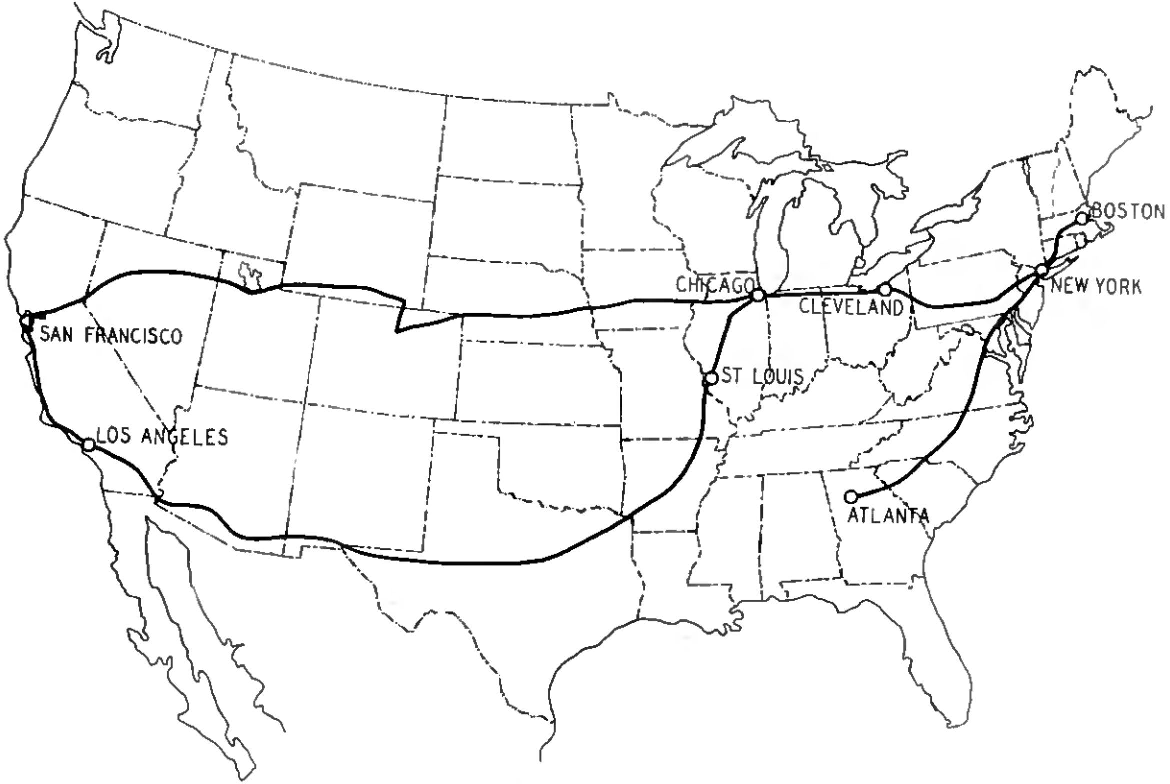


Fig. 56.—Routes of Bell System Telephoto Service in the United States.

Electrical Transmission of Pictures

A commercial service for the electrical transmission of pictures between the cities of New York, Chicago and San Francisco was inaugurated in April, 1925. The eight cities now connected to this service and the routes of the lines used in connecting them are shown in Fig. 56. In addition, a portable transmitter is provided which may be moved to any desired point. At present this is located in the city of Washington, D. C.

The pictures as transmitted are of about twelve centimeters by seventeen centimeters. Any size picture, of course, can be photographed to come within these dimensions. The detail of each picture corresponds to 39.4 lines per centimeter in each direction, that is, each picture is composed effectively of about 300,000 independent elementary areas. The line time of transmission with the present commercial system is about 7 minutes.

Pictures may be sent from any of the cities shown to one or to more of the other cities which are reached by this service. Newspapers use this service for the transmission of pictures of events of national importance or where matters arise in any part of the country of large news interest. For example, pictures of the inauguration of President Hoover were sent in this way to the newspapers in San Francisco. In view of the three hours difference in time between Washington and San Francisco the pictures were published in newspapers sold at a time of day earlier than that at which the event took place.

The majority of the pictures transmitted are for business or social purposes including pictures of legal documents, advertising material to be simultaneously released at a number of separate points, pictures showing new styles in ladies' wearing apparel, personal greetings in the handwriting of the sender and finger-prints of criminals.

The Western Union and Postal Telegraph Companies now have a service in which they will accept telegraph messages for " facsimile " transmission over this picture system between those cities which the system reaches. This service has not yet been offered long enough to show how much it will be used.

Telephony in Connection with Aircraft Operation

Telephony promises to play a very important part in the practice of commercial aviation. The Bell Telephone Laboratories are carrying out a large amount of development work on all phases of telephony for this purpose. One-way receiving sets have been developed permitting an airplane pilot to receive weather reports and to determine the direction of radio beacons. Experimental radio sets suitable for two-way conversations between a moving plane and the wire telephone system have been developed and demonstrated.

Safety of airplane travel depends a great deal on the rapid accumulation and dissemination of meteorological data. An experiment on a promising method of handling such data is being carried out on an airplane route between San Francisco and Los Angeles in the State of California. At each terminal landing field and at two intermediate fields meteorologists are located. At six periods during the day each of these is rapidly connected in succession by telephone to outlying weather observation points varying in number at the different points from three to sixteen. The information thus accumulated and coordinated at each of the four landing fields is rapidly transmitted to the other three fields by means of printer telegraph circuits connecting them. This constant rapid observation of weather conditions along the airplane route and over a considerable territory around it permits very accurate prediction of the weather conditions which any plane will meet in its travel over the route. Such weather predictions may be communicated to the airplanes before starting or by radio during their flight.

Printer telegraph circuits appear to be a particularly convenient means of interchanging information among important landing fields along airplane routes.

Ship-to-Shore Telephony

The Bell System development work on ship-to-shore telephony was originally started with wave-lengths in the neighborhood of 400 meters, which were later taken into the broadcasting range. In 1920 shore transmitting and receiving stations in northern New Jersey were equipped to operate simultaneously three separate telephone channels in this range. Through these radio stations any telephone subscriber could be connected experimentally to the steamships " Gloucester " and " Ontario " which were engaged in coastwise shipping from Boston southward. In October, 1920, a talk to one of these ships furnished an interesting part of a demonstration at a banquet in New York City tendered to the delegates to the " Preliminary International Communication Conference " which was meeting in Washington at that time.

Development of ship-to-shore telephony has been delayed because of uncertainties regarding the commercial situation and wave-length assignments. At the present time the work is again being actively pushed using wave-lengths under 100 meters. A transmitting and a receiving station will shortly be in course of construction near the seacoast of northern New Jersey and a radio-telephone set is about to be installed on the steamship "Leviathan" to operate with these shore stations. As this ship approaches or leaves New York it is expected to be possible to talk from it to any telephone in the Bell System. This is intended not only as a demonstration of the technical features of such a service but to afford an indication of the extent to which such a service will be used under commercial conditions.

Radiotelephony is being used from shore stations to coastal boats in a number of cases in the United States, but not connected to the commercial telephone system. These include particularly certain boats of the U. S. Coast Guard Service. A careful study, including tests, has been made of telephone service to tugboats operating in New York harbor for the purpose of controlling and thus making more efficient the operation of such craft. So far, it is not clear that this service will be commercially justified.

Telephony to Other Mobile Stations

Consideration has been given to telephone connections for types of mobile stations other than ships and airplanes. Communication with moving trains can technically be carried out with facilities now available. Active studies are under way to determine the practicability of providing such service at a cost which would be attractive commercially and with apparatus which can be limited to a reasonable space on the train.

Telephone Services of Railroads and Other Public Utilities

The operation of railway systems requires a large amount of communication service. The dispatching of trains was, until recent years, carried out largely by the use of telegraph. This has been rapidly changed until at the present time on over 60 per cent of the total railway mileage the train dispatching is by telephone. The railroads' telephone service to stations in the Bell System is through P.B.X.'s leased to them by the telephone companies. In addition to this, the railroads frequently own private telephone circuits extending along their rights of way which connect to and are switched through these same P.B.X.'s.

Similar arrangements are provided for meeting the special communication needs of electric power companies, oil pipe-line companies, and other utilities.

Telephone Public Address Systems

Experience in many cases has shown that with the public address system used by the Bell Companies it is possible to amplify speech or other sounds so that they can be heard by an audience of practically unlimited size. Such public address systems as they are called are used very extensively in large auditoriums and at large public gatherings. For example, the ceremonies of inauguration of President Hoover held on the steps of the Capitol in Washington were amplified by the public address system so as to be heard by a gathering estimated at a hundred thousand persons, gathered within a radius of about 300 meters.

Furthermore, by using the public address system with suitable long distance telephone circuits, it is possible to convey the proceedings of such occasions simultaneously to audiences in all parts of the country. The local distribution of such proceedings is, however, now done largely by radio broadcast rather than by use of the public address system.

A use of the public address system which so far has been taken advantage of only on a few special occasions is by providing two-way operation to interconnect two or more meetings held simultaneously in different places. A notable example of this usage is the joint meeting of the American Institute of Electrical Engineers and the Institution of Electrical Engineers in London on February 16, 1928, interconnected by the transatlantic telephone circuit. In this meeting, addresses were heard by both audiences and a resolution made in London and seconded in New York was jointly and unanimously carried. It is possible that this may foreshadow a future important use of a public address system.

Television

The possibility of transmitting pictures of a scene over electrical circuits at so high a speed that the effect is given of seeing at a distance has naturally interested telephone people for a considerable while. However, the large amount of detail which is taken in by the human eye and the resulting broad band of frequencies required to transmit this detail as well as the necessary complexity of the terminal apparatus has, so far, prevented the development of a practical service of this kind.

In 1927 the Bell engineers demonstrated to a large number of interested people a television circuit which extended from New York to Washington, a distance of about 440 kilometers. The television pictures so demonstrated had a detail corresponding to 50 lines in each direction, that is 2,500 elementary areas and 18 such pictures were shown each second. Two circuits especially corrected for volume and phase distortion over a band width of about 20,000 cycles were employed between the two cities. These circuits were, for the most part, in open wire although approximately 13 kilometers of specially loaded cable were necessary at the ends in entering the cities. By means of a separate talking circuit a person at one end of the system could talk to, as well as see, a person at the other end. Systems of approximately twice the detail and also systems adapted to the viewing of larger scenes such as athletes in action have since been developed and demonstrated.

Time Service

Arrangements have been made in many parts of the country to furnish subscribers who desire it, accurate information as to the time of day. A subscriber wishing the information asks for or dials a particular number assigned for this purpose and is connected either to an operator who advises him individually as to the time or is switched across a bus-bar to which is connected the amplified speech of an operator repeating at fifteen second intervals the exact time of day. In the present development of this service it is the practice to localize in one place the time service for an entire exchange area.

By-Products

Certain interesting and important by-products of the telephone development work justify a brief mention. Three arts separate from the telephone art have been radically changed by such by-product developments. These include submarine telegraphy, phonographs and motion pictures.

The changes in submarine telegraphy have resulted from development by the Bell Laboratories of the materials known as " permalloy " and " perminvar " which have unusual magnetic properties at low flux densities. Submarine cables so loaded can transmit approximately 10 times as many words per minute in one direction as compared to cables of the same weight as previously constructed. As such loaded cables are not duplexed the effective increase in speed of transmission is approximately five times.

Development work in connection with the faithful recording and reproduction of sound has greatly improved phonographs and their records. The "Orthophonic Victrola" is an example of such development.

An extension of this work led to the development of the "talking" motion picture. The systems known under the names " Vitaphone " and "Movietone" followed from this work. Great interest has been aroused in such systems in the amusement field in the United States.

Moving picture houses in the important cities and towns are already equipped to show pictures of this type and it appears destined to revolutionize the motion picture art.

A study of speech and hearing in connection with telephone service has led to the development of various devices of value to those having abnormal hearing or speech. This work has been carried out in close cooperation with interested members of the medical profession. One of these devices, the "audiometer," is useful in determining the condition of hearing of individuals by determining the smallest volume of sound at a considerable number of different frequencies which the individual can hear. This device, in rapidly testing large groups of people such as in the public schools, is believed to be of considerable importance. Sound amplifying devices are provided for those hard of hearing.

Another interesting by-product is an artificial larynx for those who have lost their natural larynx as a result of pathological conditions. Apparatus has also been constructed to permit the totally deaf to understand speech sounds by holding their fingers against a moving diaphragm. In one form the individual fingers and thumb are held against separate vibrating bodies and the important range of speech sounds is divided by electrical filters and one part of it applied to each of these five vibrating bodies. This partial electrical analysis of sound appears to be of considerable help in this tactual appreciation of sound.

Other tools of interest to the medical profession include electrical stethescopes and electro-cardiographs. The first of these permits any desired number to listen to chest or other sounds in medical patients. Electrical filters may be interposed in such arrangements to exaggerate or subordinate certain part of the sound. The electrocardiograph, by permitting the amplification and recording of slight differences of electrical potential between selected points of the skin of a patient give an indication of the condition of his heart beat.

Conclusion

In the above discussion, while emphasis has been placed upon engineering matters, it has naturally been impossible in the discussion of results to separate engineering considerations from many other important phases of the telephone communication problem. While engineering is essential to the results that have been obtained, they are due also to these other factors, commercial and general in their character, and to the policies as regards service and operations which guide the Bell Telephone System. Furthermore, the solution worked out has been designed specifically to meet conditions in the United States, conditions which in many respects are different in the different countries.

It is, of course, not possible in a paper of such broad scope to give technical details of the engineering problems involved. These have, however, been quite fully set forth in numerous articles in the technical press of the United States. For the convenience of those who may wish to refer further to these matters, a bibliography containing a selected list of some of the more important articles is attached to this paper.

In looking forward, there seems to be no doubt that the development of telephone communication in the United States, commercially and technically, will be more rapid than in the past, not less rapid. There are strong indications that in the future very much larger amounts of telephone service, both exchange and toll, will be demanded than at the present time, and in fact that for a number of years at least the rate of growth will continue to increase. The type and extent of services supplied will be modified to meet the broadening and multiplying demands of the changing business and social structure of the country. Finally, it is evident that the rapid advance of science will continue to bring forward new possibilities by means of which new and improved forms of communication systems, apparatus, and materials, can be developed.

These facts all indicate that the engineering work for the telephone communication system of the United States is not complete nor decreasing in magnitude or importance, but on the contrary it is increasing in volume and complexity and in the importance of the problems to be undertaken and solved.

Authors' Note

The authors wish to acknowledge their indebtedness to a large number of members of the organization for their assistance in the preparation of this paper. It is impracticable to mention all who have been of assistance but they wish to express their appreciation particularly to Messrs. O. B. Blackwell, W. E. Farnham, W. H. Harden, H. S. Osborne, and W. A. Stevens.

PARTIAL BIBLIOGRAPHY OF PAPERS RELATING TO THE BELL COMMUNICATION SYSTEM

General

Ideals of the Telephone Service. Bell Telephone Quarterly, Vol. 1, Oct. 1922, pages 1–11. Science, Vol. 57, Feb. 23, 1923, pages 219–224, Annual Report of Smithsonian Institution, 1922, pages 533–540.	J. J. Carty
Semi-Centennial of the Telephone. Bell Telephone Quarterly, Vol. 5, Jan. 1926, pages 1–11. Telegraph and Telephone Age, Vol. 44, March 1, 1926, pages 98–101.	J. J. Carty
Fifty Years of Telephone Progress, 1876–1926. Telegraph and Telephone Age, Vol. 44, Feb. 1, 1926, pages 51–53.	J. J. Carty
Building for Service. Bell Telephone Quarterly, Vol. 7, April 1928, pages 69–81.	H. P. Charlesworth
General Engineering Problems of the Bell System. Electrical Communication, Vol. 4, Oct. 1925, pages 111–125. Bell System Technical Journal, Vol. 4, Oct. 1925, pages 515–541.	H. P. Charlesworth
Bell System Research Laboratories. Electrical Communication, Vol. 2, Jan. 1924, pages 153–163.	E. B. Craft
Development and Research in the Bell System. Bell Telephone Quarterly, Vol. 4, Oct. 1925, pages 266–280.	E. B. Craft
The Budget Plan of the Bell System. Bell Telephone Quarterly, Jan. 1923, pages 32–42. Electrical Communication, April 1923, pages 64–68.	C. A. Heiss
Service in the Making. Bell Telephone Quarterly, Vol. 1, Oct. 1922, pages 26–33.	K. W. Waterson
Functions and Management Problems of the Traffic Department. Bell Telephone Quarterly, Vol. 5, Oct. 1926, pages 203–218.	K. W. Waterson
Standardization in the Bell System. Bell Telephone Quarterly, Vol. 8, Jan. 1929, pages 9–24, and April 1929, pages 132–152.	H. S. Osborne

Local Service

General

Selection of Central Office Names. Bell Telephone Quarterly, Vol. 6, Oct. 1927, pages 231–237.	A. E. Van Hagan
The Planning of Telephone Exchange Plants. American Institute of Electrical Engineers, Transactions, July 1928, pages 809–817.	W. B. Stephenson

Cable Plant

Development of Cables Used in the Bell System. Bell Telephone Quarterly, Vol. 2, Apr. 1923, pages 94–106.	F. L. Rhodes
1800-Pair Cable Becomes a Bell System Standard. Bell Telephone Quarterly, Vol. 8, Jan. 1929, pages 25–29.	F. L. Rhodes

Switching Systems

Machine Switching Telephone System for Large Metropolitan Areas. Bell System Technical Journal, Vol. 2, Apr. 1923, pages 53–89. American Institute of Electrical Engineers, Transactions, Vol. 42, Feb. 1923, pages 187–201.	E. B. Craft L. F. Morehouse H. P. Charlesworth
Machine Switching Private Branch Exchanges and Their Application to Railroad Service. In American Railway Association, Telegraph and Telephone Section, Papers, 1924, pages 418–440.	W. H. Harrison
Panel Type Machine Switching System in the United States. Electrical Communication, Vol. 4, Oct. 1925, pages 91–97.	H. P. Clausen
Telephone Switchboard—Fifty Years of History. Bell Telephone Quarterly, Vol. 7, July 1928, pages 149–165.	F. B. Jewett

Buildings

Housing the Bell System. Bell Telephone Quarterly, Vol. 5, July 1926, pages 131–139. Post Office Electrical Engineers' Journal, Vol. 19, Jan. 1927, pages 325–334.	H. P. Charlesworth

Toll Service

Short Distance Toll Service

Tandem System of Handling Short-Haul Toll Calls. American Institute of Electrical Engineers, Transactions, Jan. 1928, pages 9–20.	F. O. Wheelock E. Jacobsen

Long Distance Service

General

Engineering the Long Lines. Bell Telephone Quarterly, Vol. 2, Jan. 1923, pages 18–31.	J. J. Pilliod
Advance Planning of the Telephone Toll Plant. American Institute of Electrical Engineers, Transactions, Vol. 47, Jan. 1928, pages 1–8.	J. N. Chamberlin

Telephone Toll Lines

Telephone Transmission Over Long Distances. Electrical Communication, Vol. 2, Oct. 1923, pages 81–94. American Institute of Electrical Engineers, Transactions, Vol. 42, Oct. 1923, pages 984–995.	H. S. Osborne
Some Very Long Telephone Circuits of the Bell System. Bell System Technical Journal, Vol. 3, July 1924, pages 495–507.	H. H. Nance
Transmission Features of Transcontinental Telephony. American Institute of Electrical Engineers, Transactions, Vol. 45, Sept. 1926, pages 1159–1167.	H. H. Nance

Open Wire and Carrier Circuits

Carrier Current Telephony and Telegraphy. American Institute of Electrical Engineers, Transactions, Vol. 40, Feb. 1921, pages 205–300. Electrician, Vol. 36, May 6, 1921, pages 551–554.	E. H. Colpitts O. B. Blackwell
Practical Application of Carrier Telephone and Telegraph in the Bell System. Bell System Technical Journal, Vol. 2, Apr. 1923, pages 41–52.	A. F. Rose

Making the Most of the Line. Electrical Communication, Vol. 3, July 1924, pages 8–21.	F. B. Jewett
Carrier Systems on Long Distance Telephone Lines. Bell System Technical Journal, Vol. 7, July 1928, pages 564–629. American Institute of Electrical Engineers, Transactions, Vol. 47, Oct. 1928, pages 1360–1386.	H. A. Affel C. S. Demarest C. W. Green
Carrier Telephone System for Short Toll Circuits. American Institute of Electrical Engineers, Transactions, Vol. 48, Jan. 1929, pages 117–139.	H. S. Black M. L. Almquist L. M. Ilgenfritz

Toll Cables

Boston to Chicago Telephone Cable—Section of Largest and Longest Cable Line in the World Being Completed to Pittsburgh, Pa., by A. T. & T. Co. Telephony, Vol. 81, Dec. 31, 1921, pages 15–18.	R. W. King
Philadelphia-Pittsburgh Section of the New York-Chicago Cable. Bell System Technical Journal, Vol. 1, July 1922, pages 60–87. American Institute of Electrical Engineers, Transactions, Vol. 41, June 1922, pages 446–456.	J. J. Pilliod
Development of Cables Used in the Bell System. Bell Telephone Quarterly, Vol. 2, Apr. 1923, pages 94–106.	F. L. Rhodes

Toll Cables—Loading

Commercial Loading of Telephone Circuits in the Bell System. American Institute of Electrical Engineers, Transactions, Vol. 30, pt. 3, June 1911, pages 1743–1764.	Bancroft Gherardi
Commercial Loading of Telephone Cable. Electrical Communication, Vol. 4, July 1925, pages 24–39.	William Fondiller
Development and Application of Loading for Telephone Circuits. Bell System Technical Journal, Vol. 5, 1926, pages 221–281. American Institute of Electrical Engineers, Transactions, Vol. 45, Feb. 1926, pages 268–292. Electrical Communication, Vol. 4, April 1926, pages 258–276.	Thomas Shaw William Fondiller
Permalloy; the Latest Step in the Evolution of the Loading Coil. Bell Telephone Quarterly, Vol. 6, Oct. 1927, pages 239–246.	F. L. Rhodes

Toll Cables—Transmission

Telephone Transmission Over Long Cable Circuits. American Institute of Electrical Engineers, Transactions, Vol. 42, Feb. 1923, pages 86–97. Electrical Communication, Feb. 1923, pages 26–40. Bell System Technical Journal, Vol. 2, Jan. 1923, pages 67–94.	A. B. Clark
Building-Up of Sinusoidal Currents in Long Periodically Loaded Lines. Bell System Technical Journal, Vol. 3, Oct. 1924, pages 558–566.	J. R. Carson
Distortion Correction in Electrical Circuits with Constant Resistance Recurrent Networks. Bell System Technical Journal, Vol. 7, July 1928, pages 438–534.	O. J. Zobel

Toll Circuit Equipment

Telephone Repeaters. American Institute of Electrical Engineers, Transactions, Vol. 38, Oct. 1919, pages 1287–1345.	Bancroft Gherardi F. B. Jewett
Practical Application of the Telephone Repeater. The Western Society of Engineers Journal, Vol. 27, May 1922, pages 129–142.	H. S. Osborne
Telephone Repeaters. Electrical Communication, Vol. 1, Aug. 1922, pages 6–10; Nov. 1922, pages 27–36.	Bancroft Gherardi
Telephone Equipment for Long Cable Circuits. American Institute of Electrical Engineers, Transactions, Vol. 42, June 1923, pages 742–752.	C. S. Demarest
Echo Suppressors for Long Telephone Circuits. American Institute of Electrical Engineers, Transactions, Vol. 44, Apr. 1925, pages 481–490. Electrical Communication, Vol. 4, July 1925, pages 40–50.	A. B. Clark R. C. Mathes

Toll Line Construction

Poles. Bell Telephone Quarterly, Vol. 1, Oct. 1922, pages 34–44.	F. L. Rhodes
Bell System Sleet Storm Map. Bell System Technical Journal, Vol. 2, Jan. 1923, pages 114–121.	J. N. Kirk
Specializing Transportation Equipment in Order to Adapt It Most Economically to Telephone Construction and Maintenance Work. Electrical Communication, Vol. 1, Feb. 1923, pages 50–59. Bell System Technical Journal, Vol. 2, Jan. 1923, pages 47–66.	J. N. Kirk
Open Tank Creosoting Plants for Treating Chestnut Poles. Bell System Technical Journal, Vol. 4, Apr. 1925, pages 235–264. Bell Telephone Quarterly, Vol. 4, Jan. 1925, pages 132–142.	T. C. Smith
Recent Toll Cable Construction and Its Problems. Telephone Engineer, Vol. 32, Sept. 1928, pages 31–33.	H. S. Percival

Switching of Toll Circuits

Toll Switchboard No. 3. Bell System Technical Journal, Vol. 6, Jan. 1927, pages 18–26. Electrical Communication, Vol. 5, Apr. 1927, pages 255–259.	John Davidson, Jr.

Maintenance of Toll Circuits

Measuring Methods for Maintaining the Transmission Efficiency of Telephone Circuits. American Institute of Electrical Engineers, Transactions, Vol. 43, Feb. 1924, pages 423–433.	F. H. Best
Electrical Tests and Their Applications in the Maintenance of Telephone Transmission. Bell System Technical Journal, Vol. 3, July 1924, pages 353–392.	W. H. Harden
Practices in Telephone Transmission Maintenance Work. American Institute of Electrical Engineers, Transactions, Vol. 43, 1924, pages 1320–1330. Bell System Technical Journal, Vol. 4, Jan. 1925, pages 26–51.	W. H. Harden

International Connections

Connections in North America

Key West-Havana Submarine Telephone Cable System. American Institute of Electrical Engineers, Transactions, Vol. 41, Feb. 1922, pages 1–19.	W. H. Martin G. A. Anderegg B. W. Kendall

Connections to Europe

Telephoning to England. Radio Broadcast, Vol. 2, March 1923, pages 425–426.	R. W. King
Transatlantic Radio Telephony. Bell System Technical Journal, Vol. 2, Oct. 1923, pages 116–144. American Institute of Electrical Engineers, Transactions, Vol. 42, June 1923, pages 718–729.	H. D. Arnold Lloyd Espenschied
Transatlantic Radio Telephone Transmission. Bell System Technical Journal, Vol. 4, July 1925, pages 459–507. Institute of Radio Engineers, Proceedings, Vol. 14, Feb. 1926, pages 7–56.	Lloyd Espenschied C. N. Anderson Austin Bailey
Radio Telephone Developments of the Bell System. Bell Telephone Quarterly, Vol. 5, Oct. 1926, pages 219–237.	J. O. Perrine
New York-London Telephone Circuit. Bell System Technical Journal, Vol. 6, Oct. 1927, pages 736–749.	S. B. Wright H. C. Silent
Voices Across the Sea. North American Review, Vol. 224, Dec. 1927, pages 654–661.	Bancroft Gherardi
Transatlantic Telephony—The Technical Problem. American Institute of Electrical Engineers, Journal, Vol. 47, May 1928, pages 369–373. Bell System Technical Journal, Vol. 7, Apr. 1928, pages 161–167.	O. B. Blackwell
Transatlantic Telephone Service—Service and Operating Features. American Institute of Electrical Engineers, Journal, Vol. 47, Apr. 1928, pages 270–273. Bell System Technical Journal, Vol. 7, Apr. 1928, pages 187–194.	K. W. Waterson

Special Services

Telegraph Circuits

Metallic Polar-duplex Telegraph System for Long Small-gauge Cables. American Institute of Electrical Engineers, Transactions, Vol. 44, Feb. 1925, pages 316–325.	J. H. Bell R. B. Shanck D. E. Branson
Voice-Frequency Carrier Telegraph System for Cables. Electrical Communication, Vol. 3, Apr. 1925, pages 288–294.	B. P. Hamilton H. Nyquist M. B. Long W. A. Phelps

Telephone Networks for Program Transmission to Radio Broadcasting Stations

Telephone Circuits Used as an Adjunct to Radio Broadcasting. Electrical Communication, Vol. 3, Jan. 1925, pages 194–202.	H. S. Foland A. F. Rose
Telephoning Radio Programs to the Nation. Bell Telephone Quarterly, Vol. 7, Jan. 1928, pages 5–16.	L. N. Stoskopf
How Chain Broadcasting is Accomplished. Radio Broadcast, Vol. 12, June 1928, pages 65–67.	C. E. Dean

Electrical Transmission of Pictures

Transmission of Pictures Over Telephone Lines. Bell System Technical Journal, Vol. 4, Apr. 1925, pages 187–214.	H. E. Ives J. W. Horton R. D. Parker A. B. Clark

Telephone in Connection with Aircraft Operation

Airways Communication Service. Bell System Technical Journal, Vol. 7, Oct. 1928, pages 797–807. Aviation, Vol. 25, Oct. 6, 1928, pages 1090–1091, 1136, 1138, 1140, 1142, 1144, 1146.	E. B. Craft

Ship-to-Shore Telephony

Radio Extension of the Telephone System to Ships at Sea. Institute of Radio Engineers, Proceedings, Vol. 11, June 1923, pages 193–239.	H. W. Nichols L. Espenschied

Telephone Services of Railroads and Other Public Utilities

Telephone Equipment for Train Dispatching Circuits: A discussion of the Requirements, Development and Design of Latest Types of Equipment for High Grade Train Dispatching Systems Including Vacuum Tube Amplifiers and Loud Speakers. Electrical Communication, Vol. 2, Oct. 1923, pages 111–140.	W. H. Capen
Recent Developments in Telephone Train Dispatching Circuits. Railway Signaling, Vol. 17, Feb. 1924, pages 73–75; May 1924, pages 208–211; June 1924, pages 253–256; Aug. 1924, pages 320–322.	W. H. Capen

Telephone Public Address Systems

Use of Public Address System with Telephone Lines. Bell System Technical Journal, Vol. 2, Apr. 1923, pages 143–161. Electrical Communication, Vol. 1, Apr. 1923, pages 46–56. American Institute of Electrical Engineers, Transactions, Vol. 42, Feb. 1923, pages 75–85.	W. H. Martin
High Quality Transmission and Reproduction of Speech and Music. Electrical Communication, Vol. 2, Apr. 1924, pages 238–249. American Institute of Electrical Engineers, Transactions, Vol. 43, Feb. 1924, pages 384–392.	W. H. Martin Harvey Fletcher

Television

Television. American Institute of Electrical Engineers, Transactions, Vol. 46, June 1927, pages 913–917.	H. E. Ives
Production and Utilization of Television Signals. American Institute of Electrical Engineers, Transactions, Vol. 46, June 1927, pages 918–939.	Frank Gray R. C. Mathes
Synchronization of Television. American Institute of Electrical Engineers, Transactions, Vol. 46, June 1927, pages 940–945.	H. M. Stoller E. R. Horton
Wire Transmission System for Television. American Institute of Electrical Engineers, Transactions, Vol. 46, June 1927, pages 946–953.	D. K. Gannet E. I. Green
Radio Transmission System for Television. American Institute of Electrical Engineers, Transactions, Vol. 46, June 1927, pages 954–962.	E. L. Nelson

By-Products

By-Products of Telephone Research. Bell Telephone Quarterly, Vol. 7, Oct. 1928, pages 304–312.	R. W. King
Loaded Submarine Telegraph Cable. Bell System Technical Journal, Vol. 4, July 1925, pages 355–374. Electrical Communication, Vol. 4, July 1925, pages 60–70. American Institute of Electrical Engineers, Transactions, Vol. 44, June 1925, pages 882–890. Telegraph and Telephone Age, Vol. 43, Nov. 16, 1925, pages 524–525.	O. E. Buckley
Permalloy Loaded Cable. Electrical Communication, Vol. 2, Apr. 1924, pages 232–234.	F. B. Jewett
Man-made Ears for the Deaf; Why Many Deaf People Hear Normally in Noisy Places and Over the Telephone. Scientific American, Vol. 8, Nov. 1925, pages 320–321.	Harvey Fletcher
Recent Advances in Wax Recording. Bell System Technical Journal, Vol. 8, Jan. 1929, pages 159–172.	H. A. Frederick
Sound Recording with the Light Valve. Bell System Technical Journal, Vol. 8, Jan. 1929, pages 173–183.	D. MacKenzie
Synchronization and Speed Control of Synchronized Sound Pictures. Bell System Technical Journal, Vol. 8, Jan. 1929, pages 184–195.	H. M. Stoller
A Sound Projector System for Use in Motion Picture Theatres. Bell System Technical Journal, Vol. 8, Jan. 1929, pages 196–208.	E. O. Scriven

Miscellaneous

Telephone Transmission. Sibley Journal of Engineering, Vol. 31, Apr. 1917, pages 177–180.	Bancroft Gherardi
Transmission Unit and Telephone Transmission Reference Systems. Bell System Technical Journal, Vol. 3, July 1924, pages 400–408. American Institute of Electrical Engineers, Transactions, Vol. 43, June 1924, pages 797–801.	W. H. Martin

STATISTICS OF THE TELEPHONE INDUSTRY OF THE UNITED STATES

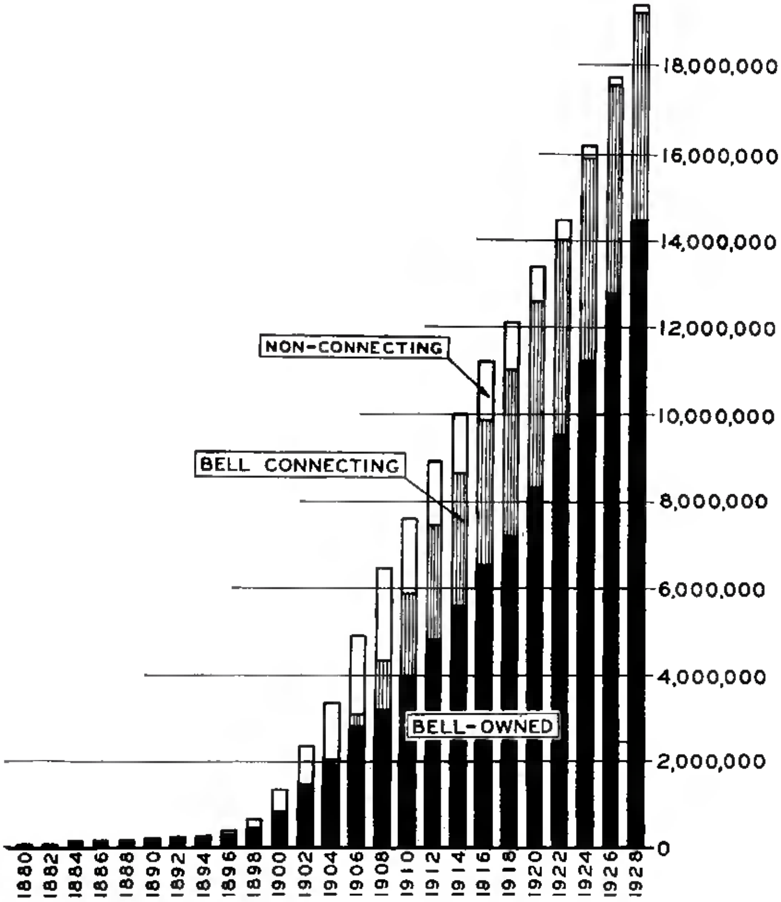


Fig. 57—Number of telephones in the United States.

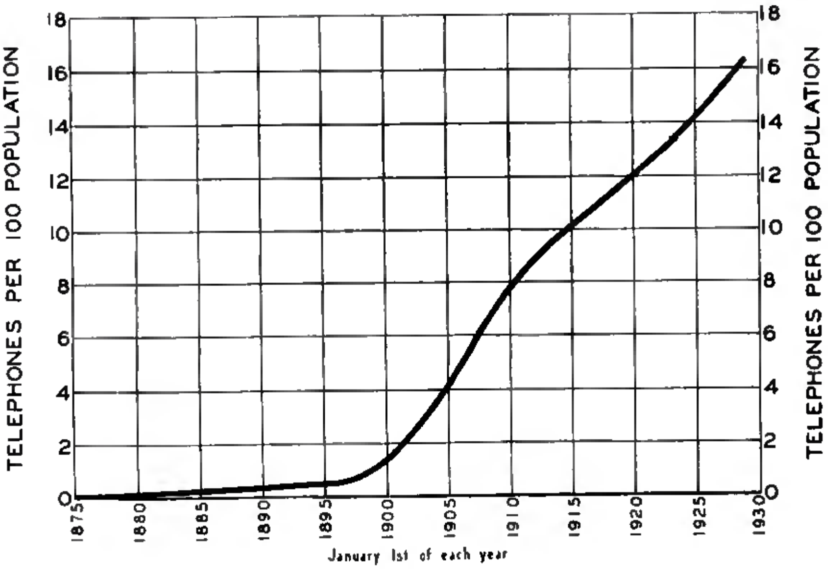


Fig. 58—Telephone development in the United States.

PERCENTAGE DISTRIBUTION OF BELL STATIONS IN FIFTEEN LARGE CITIES IN THE UNITED STATES, JANUARY 1, 1929

City	Population—Local Service Area	Number of Bell Telephones	Per cent. of Total Stations					
			Main		Private Branch Exchange	Extension	Business	Residence
			Individual	Party				
New York	6,310,100	1,702,889	51.3	5.8	34.9	8.0	53.1	46.9
Chicago	3,250,000	942,015	15.6	51.6	26.8	6.0	43.0	57.0
Philadelphia	2,040,200	375,756	25.5	40.6	24.1	9.8	46.6	53.4
Boston	1,821,400	424,781	20.8	49.4	21.4	8.4	41.2	58.8
Detroit	1,678,200	321,439	20.4	46.4	24.8	8.4	40.9	59.1
Los Angeles	1,337,000	357,504	17.4	49.5	25.2	7.9	43.6	56.4
Cleveland	1,135,800	226,186	19.0	46.9	24.4	9.7	40.1	59.9
St. Louis	1,093,500	213,041	20.5	50.6	21.3	7.6	40.9	59.1
Pittsburgh	961,000	215,125	21.9	48.6	20.4	9.1	38.7	61.3
San Francisco	751,500	252,225	27.7	31.3	31.9	9.1	48.6	51.4
Milwaukee	675,000	146,677	20.9	52.4	18.5	8.2	40.3	59.7
Washington	525,500	154,041	31.3	21.0	38.4	9.3	48.1	51.9
New Orleans	519,000	71,844	33.1	39.3	17.1	10.5	43.4	56.6
Minneapolis	492,000	126,888	30.0	40.4	20.1	9.5	33.3	66.7
Atlanta	345,000	64,546	26.8	43.8	19.4	10.0	42.7	57.3

FIG. 59

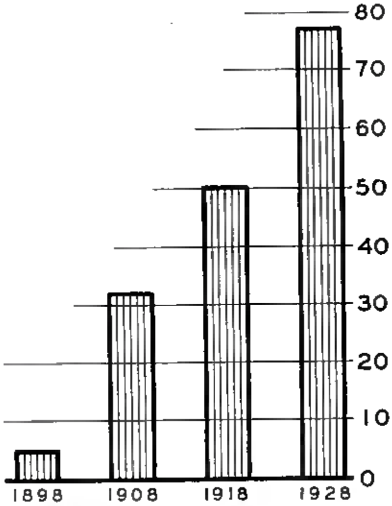


Fig. 60—Telephone conversations—Average number daily in millions in United States.

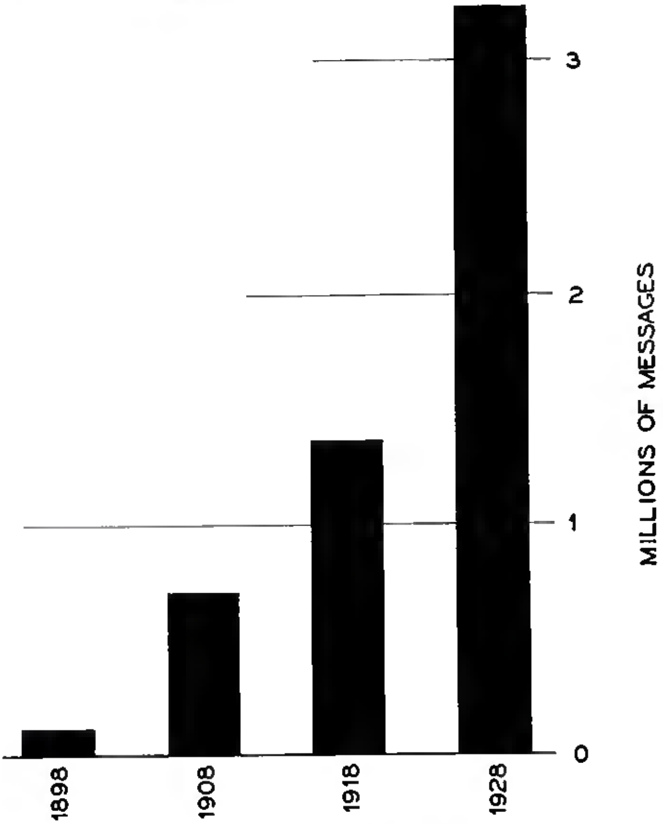


Fig. 61—Average daily number of toll messages in the United States.

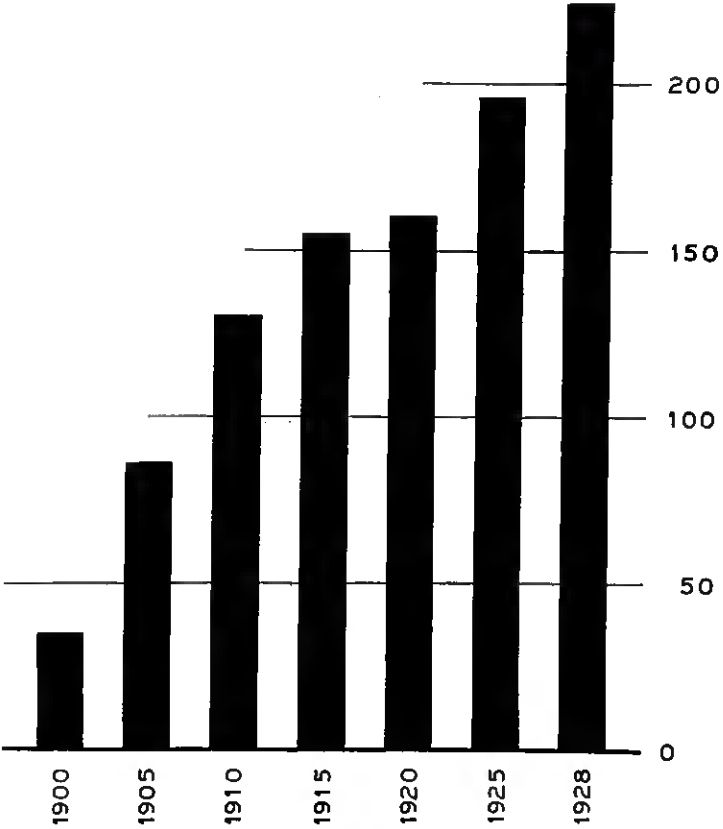


Fig. 62—Yearly telephone messages per capita in the United States.

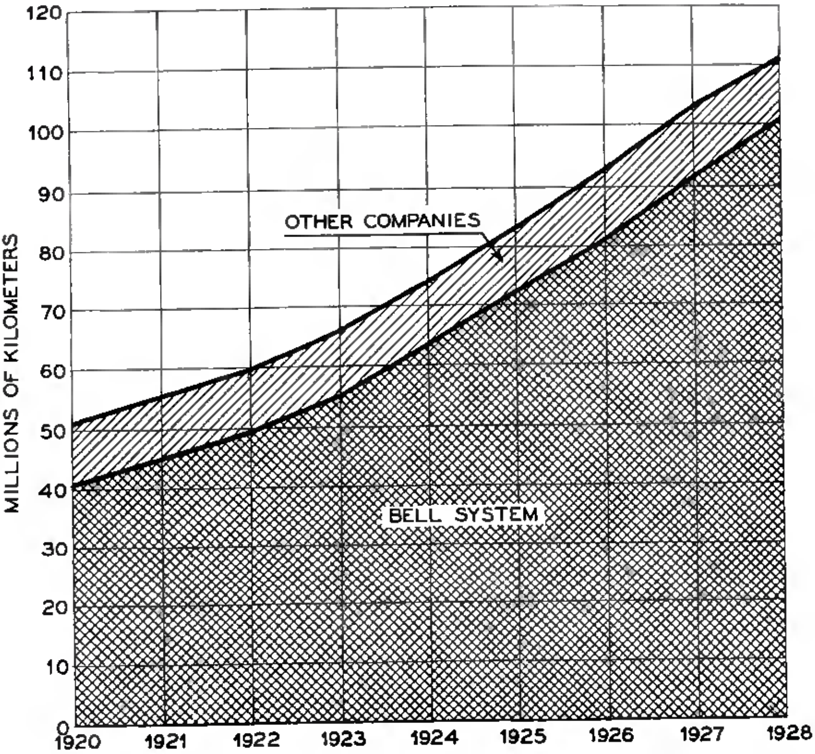


Fig. 63—Kilometers of telephone wire in the United States.

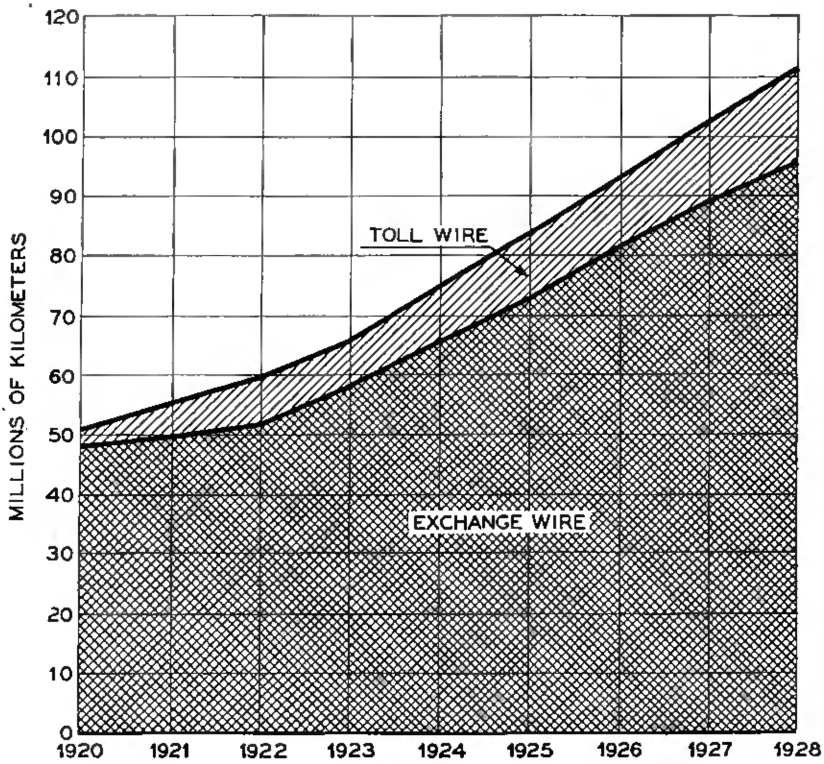


Fig. 64—Kilometers of exchange and toll wire in the United States.

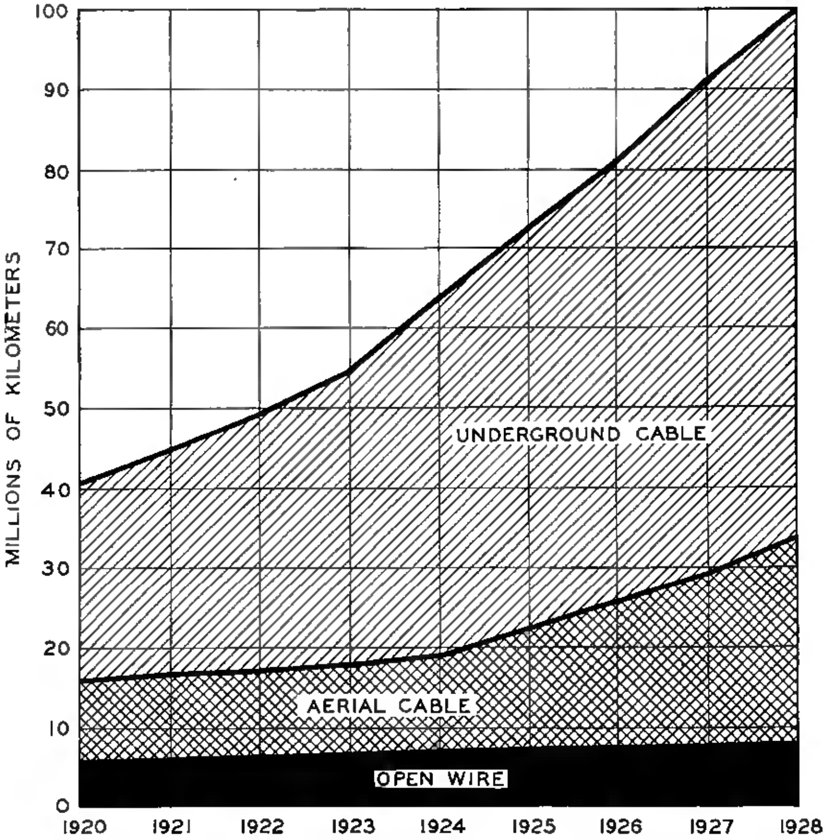


Fig. 65—Telephone wire in the Bell System in millions of kilometers.

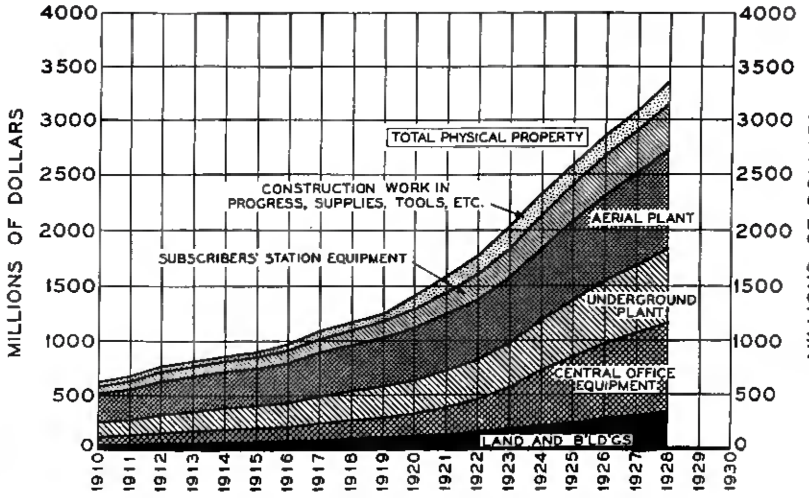


Fig. 66—Growth of various classes of physical property of the Bell System.

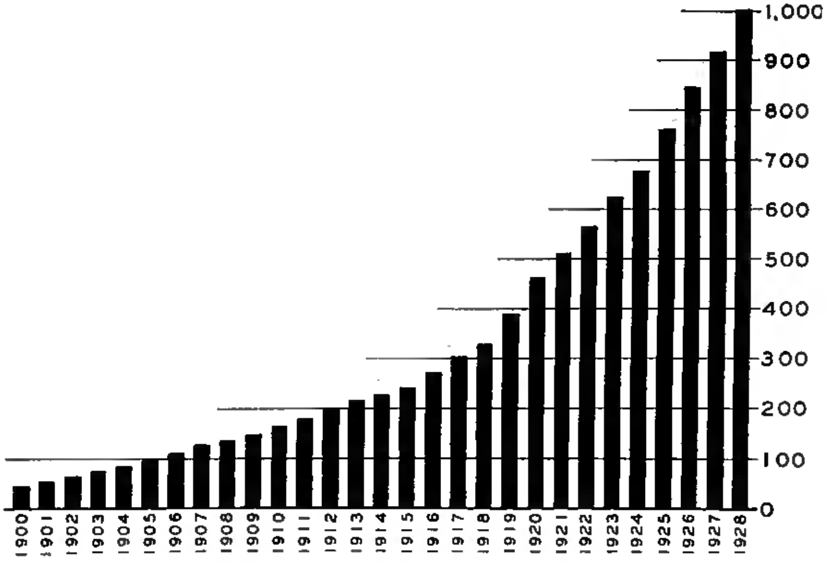


Fig. 67—Bell System revenues in millions of dollars.

TABLE SHOWING INITIAL PERIOD TOLL RATES
COMPANY, FEBRUARY 1, 1929

Air Line Kilometers (Fractions Omitted)	Station-to-Station			Person-to-Person All Hours	Report Charge
	4:30 A.M.–7:00 P.M.	7:00 P.M.–8:30 P.M.	8:30 P.M.–4:30 A.M.		
0– 19	$.10	$.10	$.10	$.20	$.10
19– 29	.15	.15	.15	.25	.10
29– 39	.20	.20	.20	.30	.10
39– 48	.25	.25	.25	.35	.10
48– 58	.30	.30	.30	.40	.10
58– 68	.35	.35	.35	.50	.10
68– 77	.40	.35	.35	.55	.10
77– 90	.45	.35	.35	.65	.10
90–103	.50	.40	.35	.70	.15
103–116	.55	.40	.35	.75	.15
116–129	.60	.45	.35	.80	.15
129–145	.65	.50	.35	.85	.15
145–161	.70	.55	.35	.90	.20
161–177	.75	.55	.40	.95	.20
177–193	.80	.60	.40	1.00	.20
193–225	.85	.65	.45	1.05	.20
225–241	.90	.70	.50	1.15	.25
241–261	.95	.75	.50	1.20	.25
261–280	1.00	.80	.55	1.25	.25
280–299	1.05	.85	.55	1.30	.25
299–319	1.10	.85	.60	1.40	.30
319–357	1.15	.90	.60	1.45	.30
357–377	1.20	1.00	.65	1.50	.30
377–396	1.25	1.00	.70	1.55	.30
396–415	1.30	1.05	.70	1.65	.35
415–454	1.35	1.10	.75	1.70	.35
454–473	1.40	1.15	.80	1.75	.35
473–492	1.45	1.20	.80	1.80	.35
492–531	1.50	1.25	.85	1.90	.40
531–550	1.55	1.30	.90	1.95	.40
550–570	1.60	1.35	.90	2.00	.40
570–608	1.65	1.40	.95	2.05	.40
608–634	1.70	1.45	1.00	2.15	.45
634–660	1.75	1.50	1.00	2.20	.45
660–685	1.80	1.55	1.05	2.25	.45
685–711	1.85	1.60	1.05	2.30	.45
711– 763	$1.95	$1.65	$1.10	$2.45	$.50
763– 814	2.05	1.75	1.15	2.55	.50
814– 866	2.15	1.80	1.20	2.75	.55
866– 969	2.35	1.95	1.30	3.00	.60
969–1072	2.55	2.10	1.40	3.25	.65
1072–1175	2.75	2.25	1.50	3.50	.70
1175–1303	3.00	2.45	1.65	3.75	.75
1303–1432	3.25	2.65	1.75	4.00	.80
1432–1561	3.50	2.80	1.90	4.50	.90
1561–1689	3.75	3.00	2.00	4.75	.95
1689–1818	4.00	3.25	2.25	5.00	1.00
1818–1947	4.25	3.50	2.50	5.25	1.00
1947–2076	4.50	3.75	2.50	5.75	1.00
2076–2204	4.75	4.00	2.75	6.00	1.00
2204–2333	5.00	4.25	3.00	6.25	1.00
2333–2397	5.25	4.25	3.00	6.50	1.00
2397–2462	5.50	4.25	3.00	7.00	1.00
2462–2590	5.75	4.50	3.25	7.25	1.00
2590–2719	6.00	4.75	3.50	7.50	1.00
2719–2848	6.25	5.00	3.50	7.75	1.00
2848–2977	6.50	5.25	3.75	8.25	1.00
2977–3105	6.75	5.50	4.00	8.50	1.00
3105–3234	7.00	5.50	4.00	8.75	1.00
3234–3363	7.25	5.75	4.25	9.00	1.00
3363–3492	7.50	6.00	4.50	9.50	1.00
3492–3620	7.75	6.25	4.50	9.75	1.00
3620–3749	8.00	6.50	4.75	10.00	1.00
3749–3878	8.25	6.75	5.00	10.25	1.00
3878–4006	8.50	6.75	5.00	10.75	1.00
4006–4135	8.75	7.00	5.25	11.00	1.00
4135–4264	9.00	7.25	5.50	11.25	1.00
4264–4393	9.25	7.50	5.50	11.50	1.00
4393–4521	9.50	7.50	5.75	12.00	1.00
4521–4650	9.75	7.75	5.75	12.25	1.00
4650–4779	10.00	8.00	6.00	12.50	1.00
4779–4907	10.25	8.25	6.25	12.75	1.00

FIG. 68

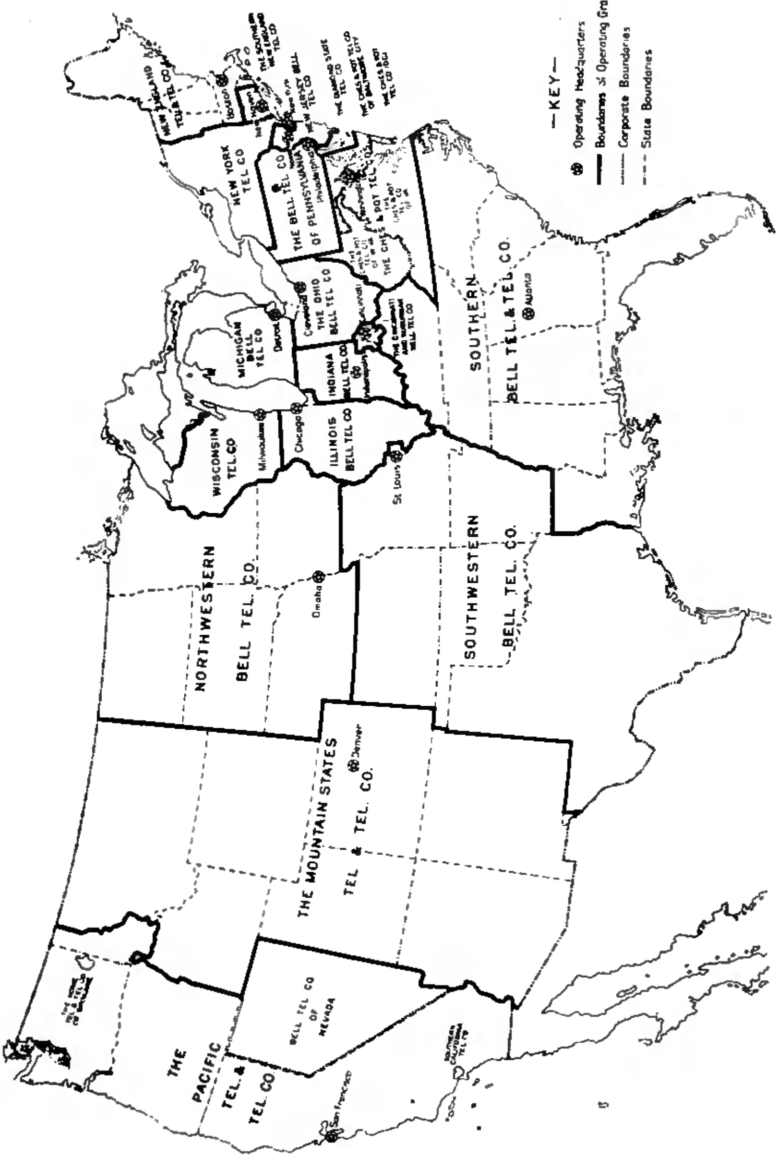


Fig. 69—The Bell Telephone System showing territories of the associated operating companies.

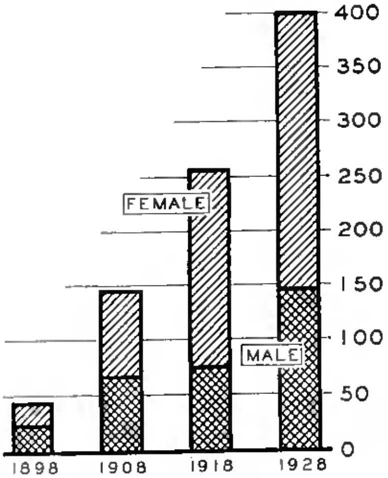


Fig. 70—Thousands of telephone employees in the United States.

TRANSATLANTIC TELEPHONE SERVICE

List of places which may be connected with the transatlantic telephone service at the present time. The figures shown both for population and telephones are estimates for January 1, 1929.

	Total Population	Total Telephones	Number Served by Transatlantic Connection	
			Population	Telephones
England, Scotland, Wales, and Northern Ireland	45,830,000	1,780,000	45,830,000	1,780,000
Germany	64,860,000	3,000,000	64,860,000	3,000,000
Belgium	8,000,000	225,000	8,000,000	225,000
Holland	7,750,000	250,000	7,750,000	250,000
Switzerland	3,990,000	236,000	3,990,000	236,000
France	41,370,000	1,000,000	37,900,000	935,000
Denmark	3,530,000	340,000	782,000	130,000
Norway	2,820,000	180,000	250,000	44,500
Sweden	6,150,000	484,000	760,000	169,000
Danzig	400,000	17,300	400,000	17,300
Spain	22,600,000	156,000	22,600,000	156,000
Austria	6,950,000	175,000	1,970,000	116,000
Hungary	8,620,000	134,000	1,000,000	50,000
Czechoslovakia	14,600,000	140,000	725,000	35,000
Gibraltar	17,000	500	17,000	500
Luxemburg	280,000	9,000	50,000	3,000
Total Europe	237,767,000	8,126,800	196,884,000	7,147,300
Spanish Morocco	1,000,000	600	37,000	300
Total Africa	1,000,000	600	37,000	300
United States	118,500,000	19,197,000	118,500,000	19,197,000
Canada	9,800,000	1,330,000	9,800,000	1,330,000
Mexico	15,500,000	70,000	1,500,000	31,600
Cuba	3,650,000	80,000	3,650,000	80,000
Total North America	147,450,000	20,677,000	133,450,000	20,638,600
Grand Total	386,217,000	28,804,400	330,371,000	27,786,200
World Total	1,930,000,000	32,800,000		
Percentage of Number Served to World Total	17%	85%		

FIG. 71

Structure and Nature of Troostite[1]

By FRANCIS F. LUCAS

In this paper the structure and nature of the constituent troostite (found in hardened steels) is discussed. High power metallography was first applied to this problem about six years ago and the early results were presented in an address before the Franklin Institute.

Since that time many improvements in technique have been developed which have resulted in better resolution and definition. The subject has been reviewed in the past two years and with the aid of the improvements in technique, hardened steels are found to be largely mixtures of the things which metallographers call martensite and troostite.

In small specimens of 0.90 per cent carbon tool steel hardened to C-65 on the Rockwell scale, innumerable particles of troostite are found. When these particles of troostite are examined by present high power methods the structure is clearly resolved into laminated pearlite. In certain stages of development of a troostitic nodule its structure borders on the verge of present methods of resolution.

Nodular troostite develops under favorable conditions as a globular mass. At the center is a nucleus about which the growth occurred. Radial, fan-shaped grains extend outward from the nucleus and these grains show orientation phenomena when revolved about the optical axis of the microscope.

It is believed that when martensite forms, the structure develops on the old austenitic crystallographic planes. Troostite appears not to follow the old austenitic system but seems to be a reorientation of the freshly transformed alpha iron about a nucleus which usually is an inclusion, a void, a sharp corner in a grain boundary or some other detail of structure.

The structure of troostite in various stages of its formation is illustrated by means of high power photomicrographs, many of which are shown at this Congress for the first time.

The following conclusions were reached:

Nodular troostite appears to be an aggregate of ferrite and carbide and in the very early stages of formation its structure is on the border of present methods of resolution. The condition of the ferrite and carbide in relation to each other is not stable—they tend to stratify, forming pearlite.

Troostitic nodules grow about a nucleus which may be an inclusion, a void, a corner in a grain boundary or some other detail of structure. The nodules contain fan-shaped radial grains.

The development of troostite results in a reorientation of the ferrite—seemingly without particular reference to the old austenitic crystallographic planes. Martensite does follow the old system of austenitic planes.

The small fan-shaped grains in nodular troostite may persist as small grains or they may undergo grain growth by union. It is a matter seemingly dependent upon the thermal treatment of the specimen.

IN a paper[2] presented before the Franklin Institute in the year 1924 some observations on the structure and probable nature of the constituent troostite were given. Two types of troostite were shown to occur in hardened steels depending on the mode of heat treatment. If a bar of 0.50 per cent carbon steel is given a taper heat treatment,

---

[1] Presented by the author before World Engineering Congress, Tokio, Japan, October 30, 1929.

[2] Lucas, "High Power Metallography—Some Recent Developments in Photomicrography and Metallurgical Research," *Journal of the Franklin Institute*, Vol. 201, February 1926.

troostite occurs which was described as flocculent border type for lack of a better designation. It seemed to be largely ferrite and appeared to be the means by which the excess constituent (in this case ferrite) appeared at the grain boundaries.

If a small specimen of the same steel is heated to a high temperature and quenched in oil or water depending on the circumstances of the experiment, a structure results which may be largely martensite needles with scattered particles of troostite. Sometimes relatively large areas on the prepared surface of the specimen may be almost entirely of the constituent troostite. As is well known, this condition is controlled by the rate of cooling in the quenching operation. The type of troostite found in uniformly heated and quenched specimens was defined as nodular troostite. This paper deals further with this particular constituent of hardened steel.

Since these early experiments in which high power metallography was first applied with success to the structures of hardened steel, there have been many improvements in technique. These improvements have resulted in a much higher order of resolution and it is the object of this paper to review the past work in the light of the improved methods now available and to present some new results.

To quote from the Franklin Institute paper:

". . . These nodules develop from innumerable nuclei throughout the austenite and martensite matrix. . . . The nuclei increase in number and the developing nodules become larger and larger. Irregularities in growth due to interference of nodules occur as the growing particles increase in number and size until finally the whole mass seems to be composed of nodules, some spherical in shape but many deformed due to mutual interference and to irregularities in growth. A selectivity or preference in crystal habit probably prevails for crystallographic planes since spines, branches, and interconnected crystallites may be found occasionally. In reality these are poorly formed nodules, growth in some one or more directions having been arrested."

"It is quite evident that if the entire mass of the metal passes through the nodular troostitic stage, this constituent must contain carbide or carbon in some form."

"When one of the globular-shaped crystal masses which has developed under favorable conditions of growth is sectioned in such a way as to divide the mass along a plane passing through the center; the nucleus is found at the center and fan-shaped grains extending from the center toward the outside. When freshly formed and under the highest powers of the microscope these radial grains have all of the appearance of a solid solution. The nodule must contain carbon in some form, as stratification soon takes place."

Moreover, it was shown that each of the fan-shaped grains is a separate crystalline unit for if a nodule of troostite is revolved about the optical axis of the microscope, these very small fan-shaped grains display orientation phenomena in exactly the same way as a system of polyhedral grains in a pure metal will do if revolved about the optical axis of a microscope while being kept under observation at 100 or 200 diameters magnification. The only difference lies in the fact that in nodular troostite the grains are fan-shaped and quite small, making it desirable to carry out the observations with an oil immersion lens which will yield high magnifications.

Fig. 1 is reproduced from the Franklin Institute paper and illustrates a typical section on a plane passing through the center of a single nodule. Fig. 2, also from the same paper, is a diagrammatic representation of how the fan-shaped grains develop along axes of crystallization A, B, C, etc.

Fig. 1 not only shows crystallization about a nucleus but it also supplies evidence for the conclusion at that time that nodular troostite is either a solid solution of iron carbide in iron or it is a very fine aggregate of iron and iron carbide—the carbide so finely dispersed as to lose its identity under the microscope. Failure at that time to resolve the structure of troostite into its ultimate constituents compelled one to recognize the existence of the two possibilities as to structure.

The improvements in technique previously mentioned have thrown some new light on the structures found in hardened steels and these have been discussed in later papers.[3, 4, 5]

Certain it is that the structures found in hardened steel are largely mixtures of the things which metallographers call martensite and troostite, the name martensite in this case meaning a needle-like structure. Troostite, generally, is regarded as a lower order of decomposition than martensite. This, however, is not believed to be substantiated by the evidence.

It has been shown[6] that martensite is a decomposition of the austenite along the octahedral crystallographic planes. That is, martensite is a structure superimposed by decomposition of the austenite on the old crystallographic system of the austenite. Two changes are in-

---

[3] Lucas, "A Résumé of the Development and Application of High Power Metallography and the Ultra Violet Microscope," Vol. I, *Proceedings International Congress for Testing Materials*, Amsterdam, September 1927.

[4] Lucas, "Photomicrography and Its Application to Mechanical Engineering," *Mechanical Engineering*, Vol. 50, March 1928.

[5] Lucas, "Further Observations on the Microstructure of Martensite," *Trans. American Society for Steel Treating*, Vol. XV, February 1929.

[6] Lucas, "The Micro-Structure of Austenite and Martensite," *Trans. American Society for Steel Treating*, Vol. VI, No. 6, December 1924.

volved: an allotropic change of the iron from gamma to alpha and a precipitation of the carbide Fe_3C. This matter is more fully discussed in a recent paper[5] to which those interested are referred.

Troostite is not like martensite in respect to habit of formation. It does not assume fully the old austenitic crystalline symmetry. It seems to have a new crystalline orientation of its own.

Troostite develops along grain boundaries and within the grains. It may develop as a spine or branch along a crystallographic plane. The nodules may be roughly spherical masses; they may be semi-spherical masses, the flat side being bounded by a crystallographic plane or a grain boundary, or they may be rounded but irregular shaped masses. In any event, whether in ball-shaped masses or some constricted form, the small fan-shaped grains are found radiating from a nucleus of growth. This nucleus in most cases can be identified as an inclusion, a void, or a sharp corner in a grain boundary. Thus the tendency is for reorientation of the iron to occur when nodular troostite develops.

It is well known that a slow rate of cooling promotes more troostite. Rapid cooling results in more needles or the constituent we call martensite. Evidently when the rate of cooling is favorable the freshly transformed alpha iron is given time to reorient itself and does do so by growing about some convenient inclusion or other body. Thus nodular troostite develops. Whether the carbide is held in solid solution in the freshly transformed alpha iron seems to be a matter of speculation.

A small specimen of steel weighing less than ten grams, heated to 1000° C. in a vacuum for a suitable length of time, and quenched in ice and brine will contain almost innumerable troostitic bodies, many of them very small, and some quite large. The larger ones perhaps are a few ten-thousandths of an inch in diameter and from this dimension the troostitic particles decrease in size to the vanishing point of present microscopic resolution which is around 200 atom diameters.

The specimen itself will have a hardness on the Rockwell scale of about C-65. Nevertheless the troostitic bodies have been clearly resolved to show the presence of fully laminated pearlite. So that in a small specimen of steel quenched from a high temperature in a very effective cooling bath, one finds not only the needle constituent martensite but nodules of troostite containing fan-shaped grains of fully stratified pearlite.

The question naturally arises as to whether the steel in its transition from austenite to pearlite first develops a needle structure (martensitic) and then this in turn is replaced by a nodular (troostitic) one.

Some light[6] was thrown on this angle of the problem by a high power examination of an iron carbon alloy. The carbon content was 2.65

per cent and by quenching small pieces from very high temperatures, polyhedral grains of austenite containing martensitic needles and troostitic nodules were found to occur. Both constituents were found to occur in the same grain and both seemed to be entirely surrounded by austenite. Had the needles formed first and the nodules developed from the needles, one might expect to find some nodules with untransformed needles sticking out around the boundaries of the nodules. This was found not to be the case. The boundaries of the troostitic nodules are always sharply defined.

In some specimens of commercial plain carbon steels in which some tempering had taken place troostitic nodules were found in which it appeared that the nodule had grown at the expense of some martensitic needles. The needles seemed to be dimly visible in outline in the background of the nodule. Cases of this kind appear very infrequently. Microscopic evidence does not support the conclusion that one type of structure replaces the other.

If a specimen of commercial tool steel heat treated to produce some troostite in a martensitic matrix is tempered, one might expect the troostite nodules to grow in size if the nodular form of structure replaces the needle structure. As a matter of fact the nodules remain the same size and the carbide which they contain tends to coalesce into small globular particles, marking not only the border outline of the nodule but also the outlines of the fan-shaped grains.

The needle and nodular patterns are structures which result from quenching and not from tempering. The excess constituent in the case of hypo- or hyper-eutectoid steels appears to be eliminated or cleared by means of the constituent troostite. The constituent martensite (needles) appears not to be involved in this phenomena in quenched specimens when both troostite and martensite are present.

If one examines a normalized specimen of plain carbon tool steel of about 0.90 per cent carbon content he will find a large polyhedral structure marked by a carbide network, but within these grains will be found a great many smaller grains of pearlite, usually fan-shaped. In many cases the outlines of the old troostitic nodules can be traced without difficulty. From the configuration of the pattern it seems likely that these small grains within the larger (old austenitic) grain must differ in their inner crystalline symmetry, i.e., it is probable that the ferrite is not everywhere oriented the same throughout the old austenitic grain. Under some circumstances controlled by heat treatment, it appears that grain growth does occur among these small fan-shaped grains and the old austenitic grain may be uniformly oriented ferrite containing spheroidized particles of cementite which by their positions mark the old structure and tell the history of the transformations.

Professor Honda believes that martensite forms first and troostite develops secondly, replacing the martensite.[7] In his discussion of the subject he appears to deal with the ultimate nature and composition of the constituents and not with their outward form.

A number of typical illustrations are included to show in detail the structure of troostite. For these experiments a high grade tool steel of about 0.90 per cent carbon was used. Small specimens weighing about 10 grams were suitably heated in a vacuum furnace to a high temperature and quenched in ice and brine solution. The hardness of the specimens was quite uniform and averaged C-65 on the Rockwell scale. From a study of the photographs it is quite apparent that in a specimen of the kind, we may have not only the constituents martensite and troostite, but in the troostite also the constituent pearlite in the form of fan-shaped grains. From the work of Mathews,[8] Bain,[9] Enlund,[10] and others, it is also apparent that some retained austenite may be present. Hardened steel, therefore, is a complex structural aggregate at best.

Conclusions

Nodular troostite appears to be an aggregate of ferrite and carbide, and in the very early stages of formation its structure is on the border of present methods of resolution. The condition of the ferrite and carbide in relation to each other is not stable; they tend to stratify forming pearlite.

Troostitic nodules grow about a nucleus which may be an inclusion, a void, a corner in a grain boundary or some other detail of structure. The nodules contain fan-shaped radial grains.

The development of troostite results in a reorientation of the ferrite, seemingly without particular reference to the old austenitic crystallographic planes. Martensite does follow the old system of austenitic planes.

The small fan-shaped grains in nodular troostite may persist as small grains or they may undergo grain growth by union. It is a matter seemingly dependent upon the thermal treatment of the specimen.

---

[7] Honda, "Is the Direct Change from Austenite to Troostite Theoretically Possible?" *Journal British Iron and Steel Institute,* 1926, Vol. CXIV, No. 2.

[8] Mathews, "Austenite and Austenitic Steels," *Trans. American Institute of Mining and Metallurgical Engineers,* Vol. LXXI, 1925; "Retained Austenite," *British Iron and Steel Institute,* 1925, No. 11, Vol. CXII.

[9] Bain, "The Persistence of Austenite at Elevated Temperatures," *Trans. American Society for Steel Treating,* Vol. VIII, 1925.

[10] Enlund, "On the Structure of Quenched Carbon Steels," *Journal British Iron and Steel Institute,* 1925, Vol. CXI, No. 1.

Fig. 1—Mag. 3230X. Fig. 1 is reproduced from the *Journal of the Franklin Institute* and shows a typical troostitic nodule sectioned on a plane passing through the center. The nodule has developed as a globular mass about a nucleus. Radial grains have developed. These grains change from light to dark when the nodule is revolved about the optical axis of the microscope. Therefore it is clear that the small fan-shaped grains are differently oriented. Where nodular troostite forms regranulation must occur. Where changes in orientation take place, grain boundaries must result. The structure of the troostite has not been fully resolved in this photograph. Compare with others which follow.

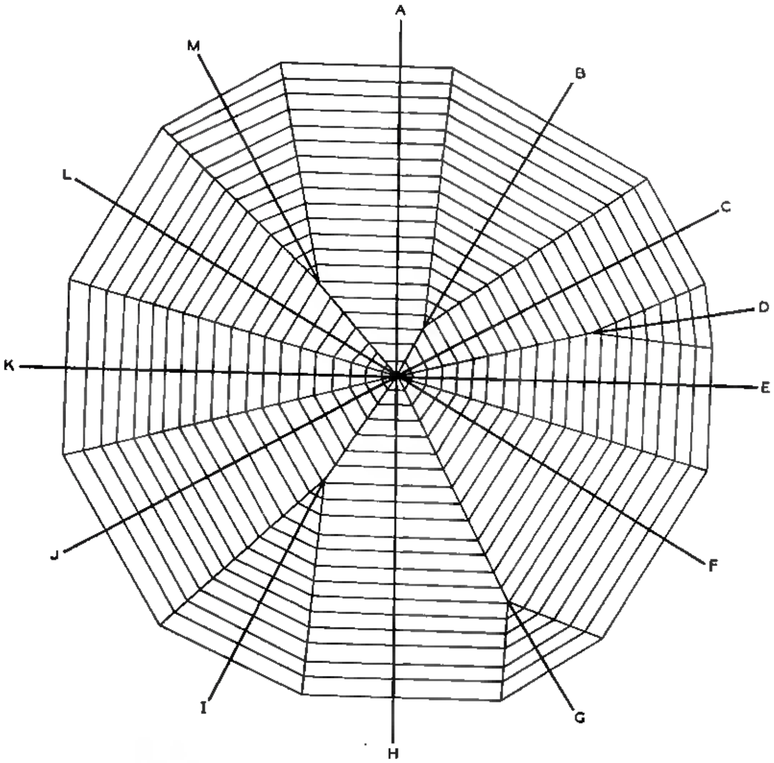


Fig. 2—A Diagram. Fig. 2, also from the *Journal of the Franklin Institute*, illustrates diagrammatically the mode of crystalline growth in a troostitic nodule.

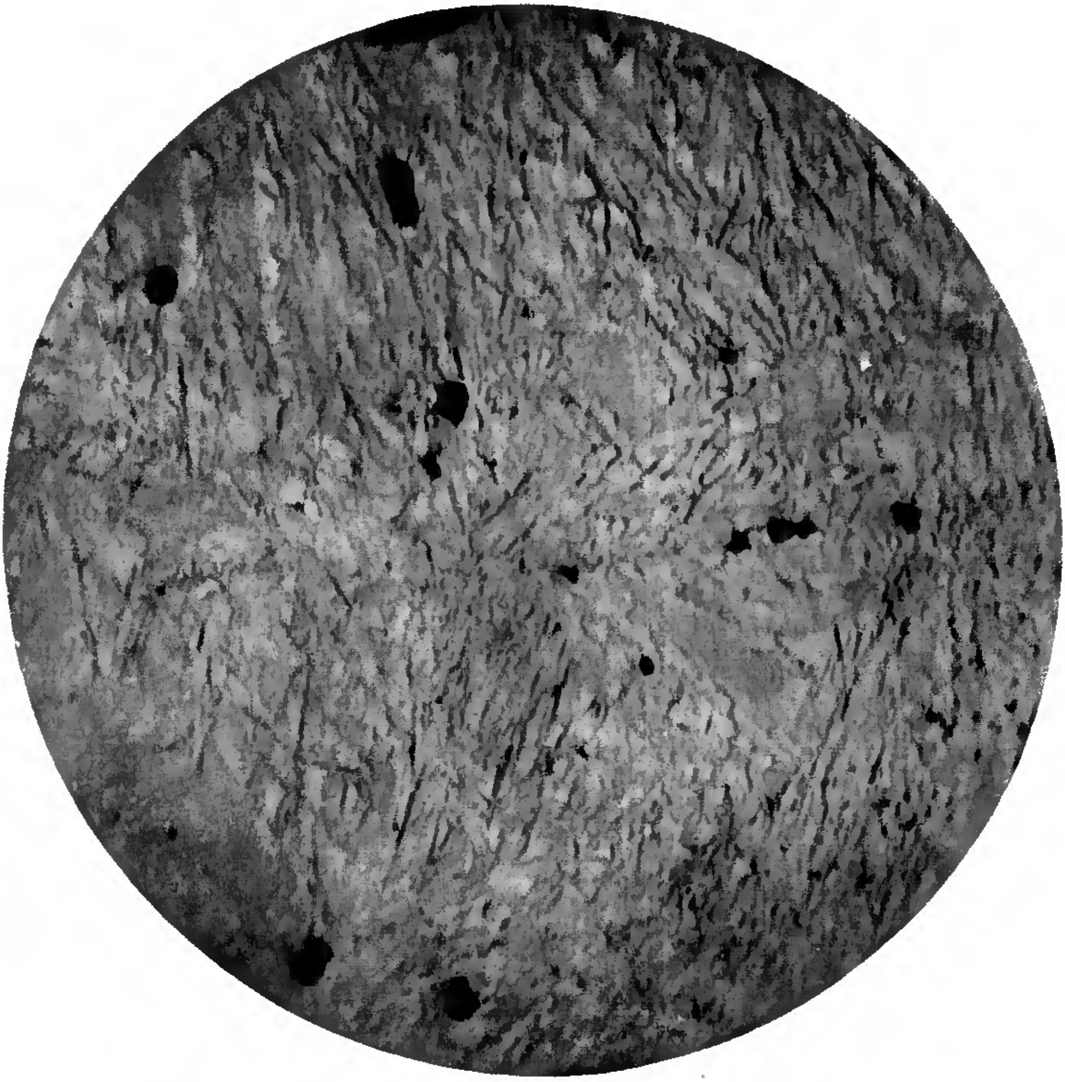


Fig. 3—Mag. 3500X. Fig. 3 shows the early stages in the formation of troostitic nodules. The background will be recognized as martensite. The dark particles are troostite. The field is on the border of an area containing large well developed nodules. This position in the specimen is one in which thermal conditions promoted the development of the needle structure but did not fully inhibit the development of nodules. The very small dark particles are about five-millionths of an inch in diameter. The larger ones are from about ten to twenty times larger.

Fig. 4—Mag. 3500X. Fig. 4 shows a somewhat later period in development of troostite. The troostite appears to have formed along grain boundaries. The excess constituent is clearly seen, and here and there a laminated structure, pearlite. Evidently whatever the state of the carbide with reference to the iron in troostite—whether contained in solid solution as first formed or whether disposed as a fine aggregate with the iron—the condition must be very unstable, otherwise evidences of finely laminated pearlite would be lacking.

Fig. 5—Mag. 3500X. Fig. 5 is of a small but well developed nodule showing four fan-shaped grains about a nucleus of growth. Some excess constituent has appeared but only the very early stages in the process of stratification to pearlite are visible. It is apparent, however, that the nodule is not composed of a solid solution. The nodule is about eight ten-thousandths of an inch in diameter.

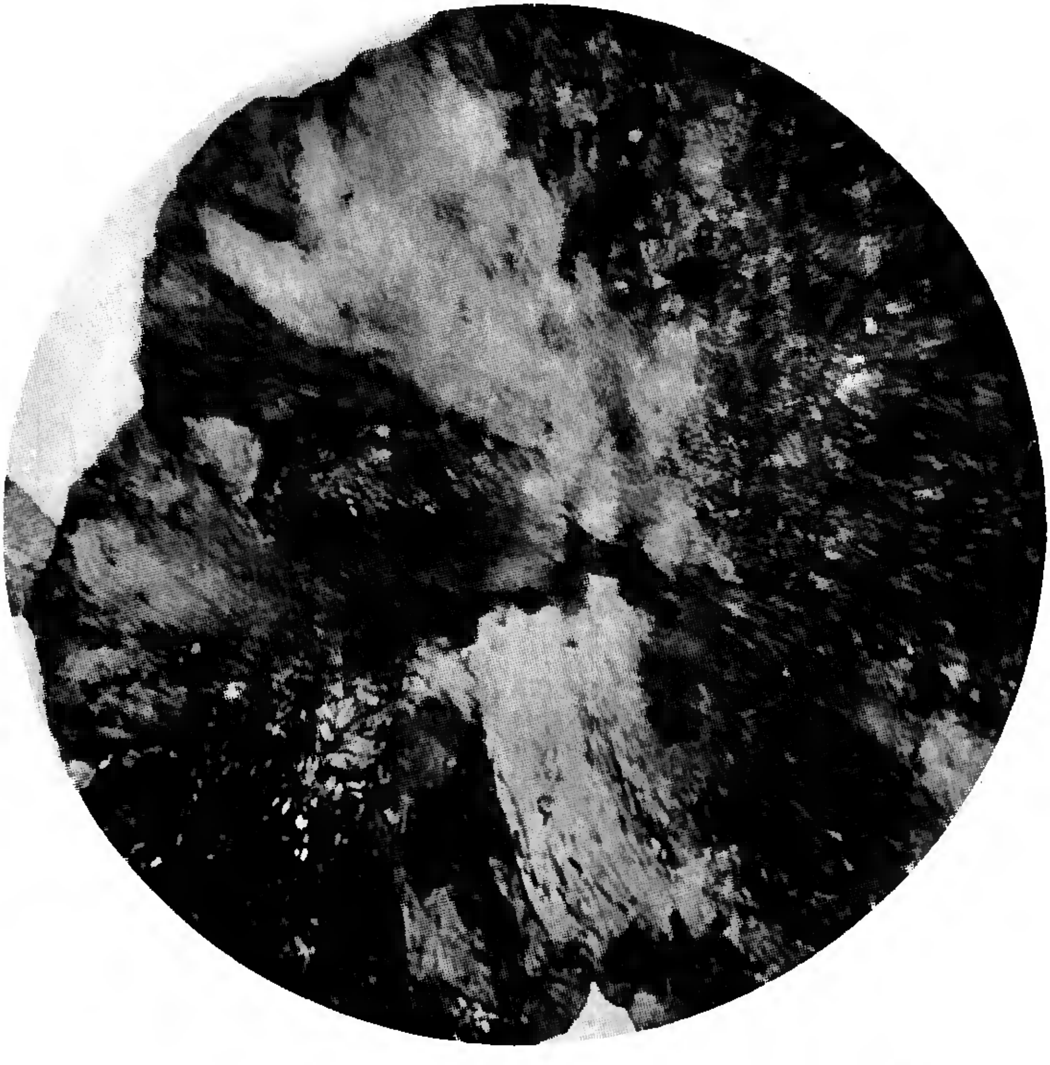


Fig. 6—Mag. 3500X. Fig. 6 is of a larger nodule than illustrated in Fig. 5 and shows obviously a little later stage in the process of stratification. This nodule is about the same as the one illustrated in Fig. 1 except that the structure has been resolved.

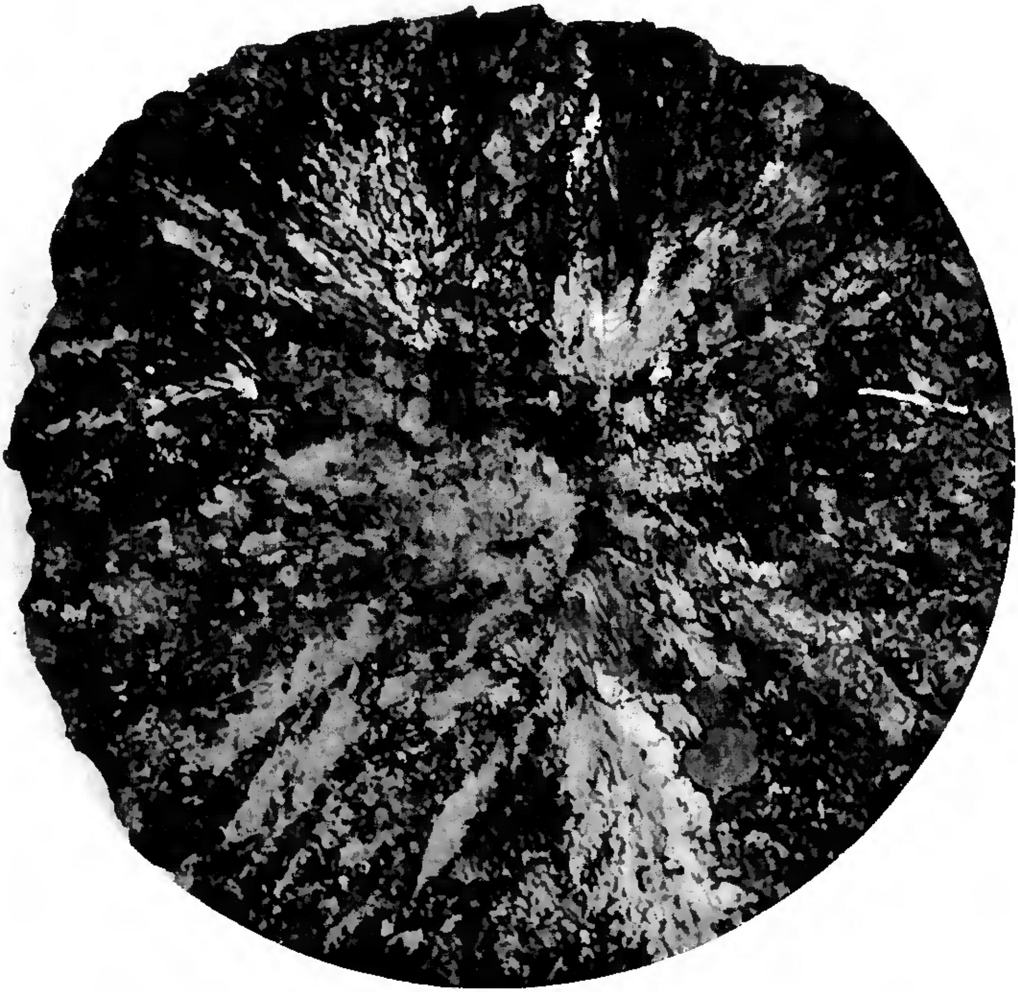
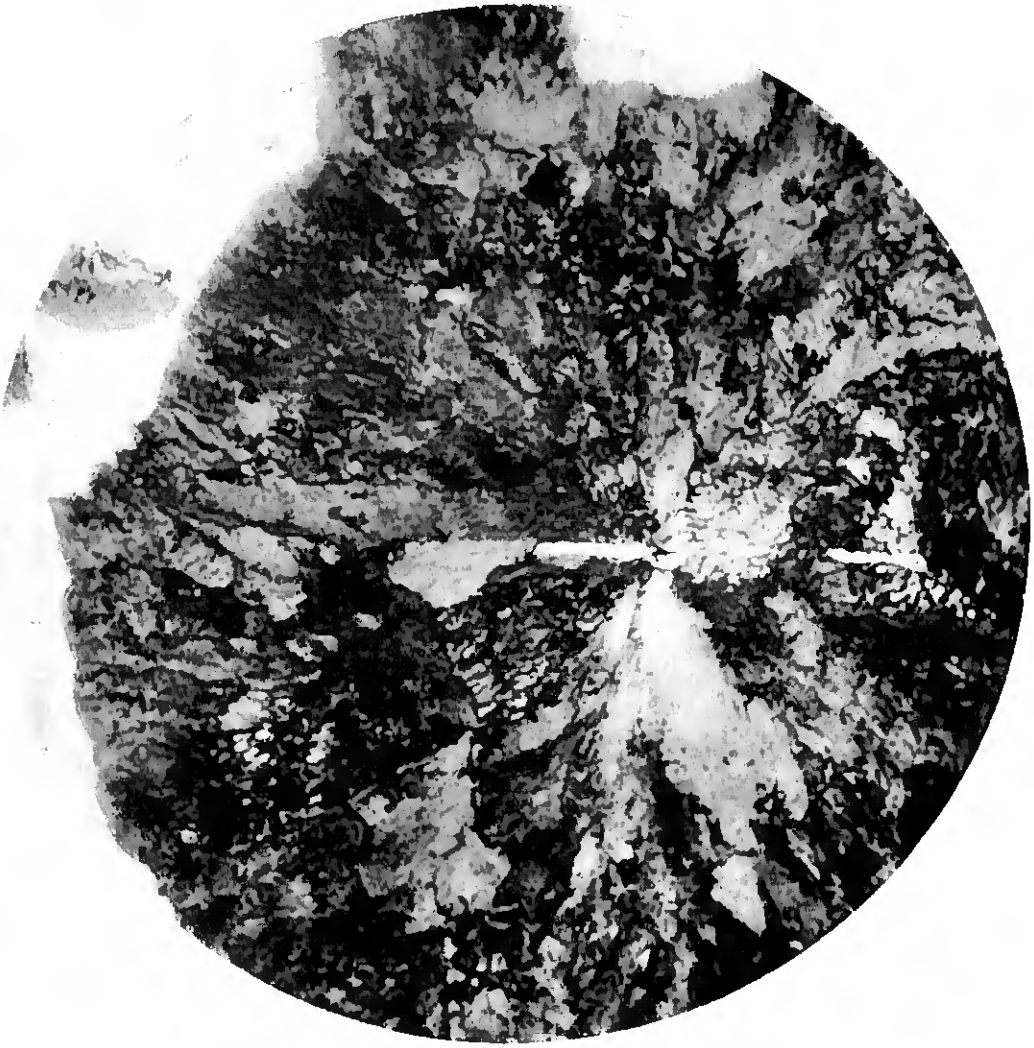


Fig. 7—Mag. 3500X. Fig. 7 illustrates a condition somewhat further advanced than that shown in Fig. 6. Stratification is well advanced and is plainly shown throughout.

Fig. 8—Mag. 3500X. Fig. 8 shows a troostitic development which had formed along an old austenitic grain boundary. The excess constituent is starting to clear at the grain boundary. Well developed pearlite is revealed. The large light-colored grain covering the center of the field is just starting to break up into two constituents. Formerly grains of this kind, because of lack of resolution, were thought to be in all probability solid solution grains. These grains must represent a state very nearly that of freshly formed troostite.



Figs. 9 and 10—Mag. 3500X. Figs. 9 and 10 show two typical nodules along grain boundaries or crystallographic planes. The excess constituent, the radial grains, some practically irresolvable and others fully resolved, and the center of growth (in Fig. 9) are clearly revealed.

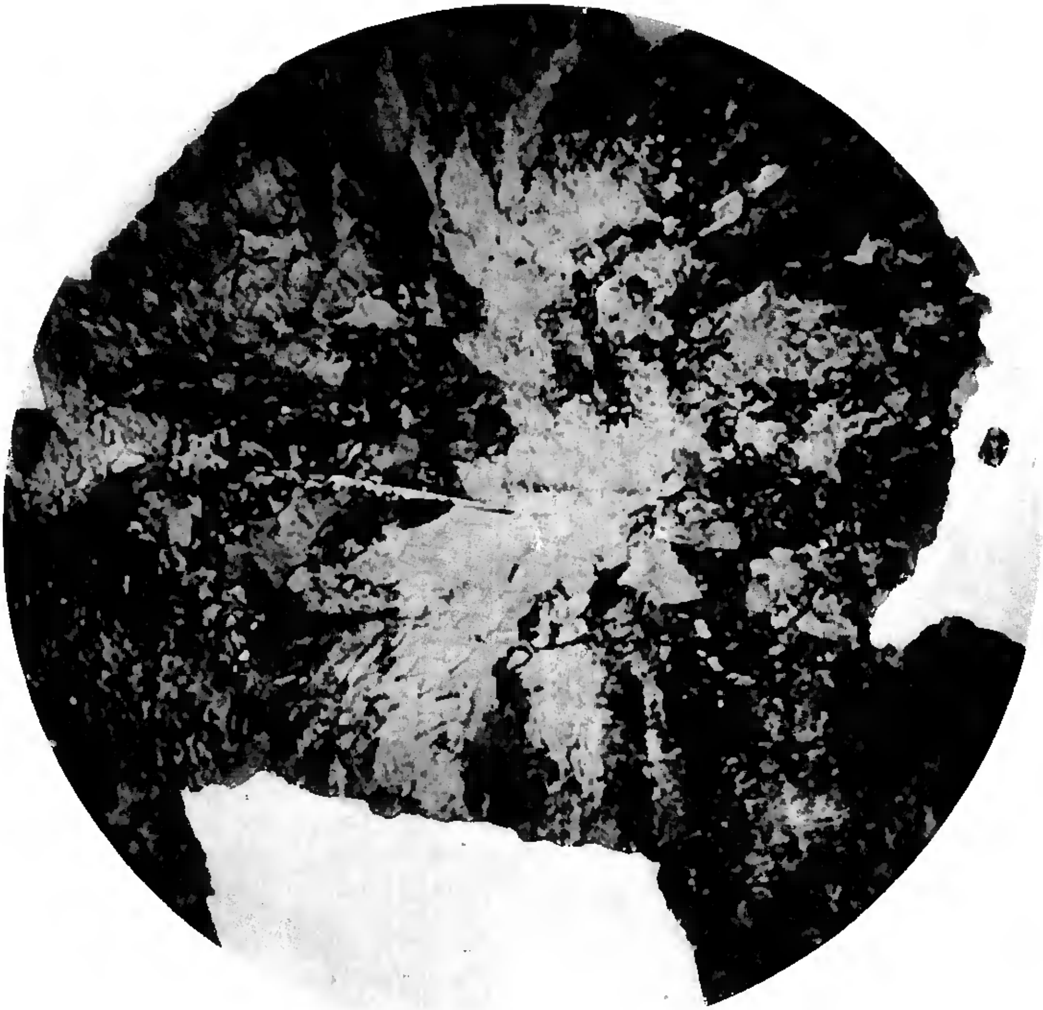


Fig. 10—Mag. 3500X.

Fig. 11—Mag. 3500X. Fig. 11 illustrates the wide range in structure to be found in hardened steel. The background is martensite which contains a troostitic nodule. One grain of the nodule is fully laminated pearlite. The other grains are in all stages of stratification.

Fig. 12—Mag. 3500X. Fig. 12 illustrates the condition which prevails when growing nodules interfere and the whole area is troostite. Small fan-shaped grains are found in the different stages of stratification.

Fig. 13—Mag. 3500X. Fig. 13 is of a field similar to Fig. 12 except that a more advanced stage in stratification is present. This photograph is reproduced from the *Proceedings of the International Congress for Testing Materials.*

Fig. 14—Mag. 3500X. Fig. 14 is of a specimen which was quenched and then drawn for ten minutes at 650° C. The outline of a troostitic nodule is clearly marked by globular carbide particles. The hardness of the specimen was C-28 after tempering.

Radio Broadcasting Transmitters and Related Transmission Phenomena[1]

By EDWARD L. NELSON

This paper is a brief discussion of recent developments in American practice concerning radio broadcasting transmitters. Descriptive material and photographs pertaining to several new commercial transmitting equipments are included. Reference is also made to the more important aspects of the related transmission problem. On account of the scope of the subject, the treatment is necessarily superficial, but it may serve to indicate the present status of the transmitter art and its relative position with respect to the industry as a whole. A short bibliography containing some of the more important recent contributions to the subject is attached as an appendix, to which reference may be had for more detailed information.

RADIO TRANSMITTERS

THE radio transmitter is essentially a focal point in the present-day broadcasting system, since upon it the program circuits converge and from it the radio distribution network emanates. For this reason, the requirements which have been imposed on transmitting apparatus are extremely rigorous, and all phases of transmitter performance have been subjected to the most careful scrutiny. Under these stimulating influences, the last few years have brought about some very noteworthy advances in this portion of the broadcasting field.

As long as music and entertainment continue to hold a prominent place on broadcasting programs, fidelity of transmission will probably remain the most sought-for characteristic, not only for the radio transmitter itself, but for all of the apparatus units in the system. A very high standard of performance has now been attained in this respect. Fig. 1, below, shows the overall frequency-response characteristic of a new type 50-kw. equipment, the first of which has gone into service at one of the leading American broadcasting stations within the past few months. It will be noted that this characteristic is substantially flat between 30 and 10,000 cycles. The greatest departure from the horizontal line which is the ideal characteristic is less than 1 db. The frequency discrimination which this represents is of such a low order that it probably could not be detected in ordinary listening tests, even by a skilled musician.

Another recognized prerequisite to a high degree of fidelity is exact proportionality between audio input and sideband output. Increased

---

[1] Read before the World Engineering Congress, Tokio, Japan, October, 1929; *Proceedings of Institute of Radio Engineers*, November, 1929.

emphasis on accurate reproduction has recently led to the introduction of improved technique for checking this important characteristic under dynamic conditions. The method employed consists of impressing a pure sine-wave input on the transmitter at various frequencies throughout the audio range and subjecting the output of a straight-line rectifier to harmonic analysis. One type of harmonic analyzer which has been used with excellent results is that due to Wegel and Moore.[2] This device produces a photographic record, an example of which is shown in Fig. 2. Measurements of this type are of particular importance under present conditions since current American practice is tending toward the extensive use of transmitters in which modulation is accomplished at relatively low power levels and the required power output is obtained by means of subsequent stages amplifying modulated radio-frequency power. Such amplifying stages

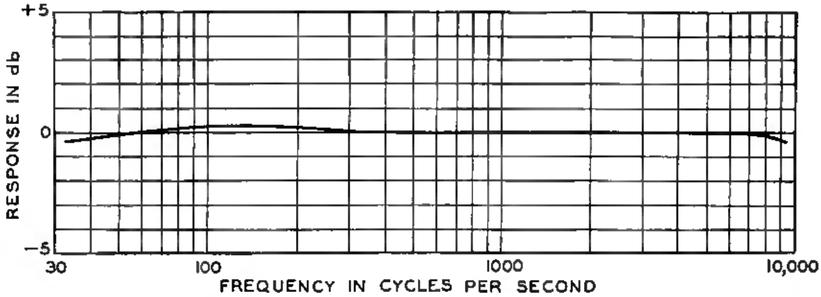


Fig. 1—Frequency-response characteristic of Western Electric 7-A (50-kw) radio transmitter.

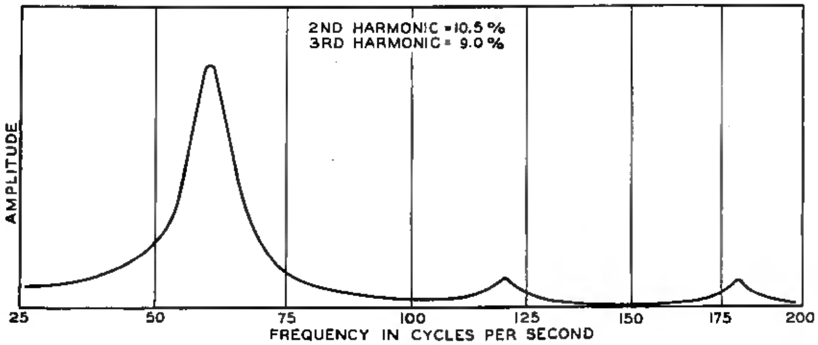


Fig. 2—Harmonic analyzer graph indicating overloading (2nd harmonic, 10.5 per cent; 3rd harmonic, 9 per cent).

---

[2] R. L. Wegel and C. R. Moore, "An Electrical Frequency Analyzer," *Bell Syst. Tech. Jour.*, p. 299–323, April, 1924.

are susceptible of serious amplitude distortion unless the conditions under which the tubes operate (direct plate and grid voltages and impedance of the connected load) are carefully predetermined. For this purpose, the harmonic analyzer has proved to be invaluable. Through its use, commercial transmitters are now available in which, at the working upper limit of modulation, the harmonics generated are not greater than 5 per cent.

The attainment of such high standards for fidelity leaves little opportunity for progress, and it is improbable that significant advances in this direction will be made in the near future. Accordingly, in continuing their efforts toward further improvements in broadcasting service, transmitter engineers have been led to divert their attention to the problem of rendering less conspicuous and objectionable the background of noise and interference which, in the past, has so seriously impaired the artistic effect of programs except in the immediate vicinity of transmitting stations. This is the principal justification for the present movement toward higher power outputs for broadcasting stations. It has also resulted in increased emphasis on the maintenance of a high average degree of modulation, a development which is rapidly bringing about a very perceptible improvement in general broadcasting conditions.

The degree of modulation of the carrier in a radio telephone transmitter is a somewhat intangible factor which necessarily varies rapidly through wide limits during the rendition of a program. With every transmitter, however, there is a definite modulation limit which is a characteristic of the design and which cannot be exceeded without bringing about serious distortion. This limit is an important performance index which, for lack of a better name, has been called "modulation capability." The modulation capability of a transmitter may be defined as the maximum degree of modulation (expressed as a percentage) that is possible without appreciable distortion, employing a single-frequency sine-wave input and using a straight-line rectifier coupled to the antenna in conjunction with an oscillograph or harmonic analyzer to indicate the character of the output.

For a number of reasons, some technical and some economic, many of the broadcasting transmitters in use have been so constructed that overloading of the audio power stage with consequent distortion occurs whenever the degree of modulation exceeds approximately 50 per cent. The usual practice in placing broadcasting transmitters into service consists of determining, by means of a suitable vacuum-tube voltmeter or other "volume indicator," the audio level at the input of the set for which distortion becomes evident. The average operating level is then established at a suitably lower value, frequently 6–10 db. Recently, by modulating at low power levels, transmitters have been produced which are capable of 100 per cent modulation without noteworthy distortion. It is obvious that, if a transmitter of this latter type is employed and the same margin is observed in determining the average audio input level, the resulting sidebands will be twice the amplitude of those produced by a transmitter whose modulation capability is only 50 per cent. To produce equivalent sidebands with a transmitter capable of but 50 per cent modulation requires that the carrier amplitude be doubled or the power output multiplied by four. In other words, insofar as signal-to-noise ratio is concerned, which is the factor that usually determines the coverage of a broadcasting station, the increase in modulation capability mentioned results in an improvement that in the older type of apparatus could only be had by quadrupling the rated output of the transmitter. From the coverage standpoint, the night range of a given station can be approximately doubled in this manner. Since this is accomplished without increase in the carrier power, the outlying zone in which the station may produce serious beatnote interference with others assigned to the same channel will not be extended. Accordingly, the use of transmitters capable of a high degree of modulation is a notable contribution toward the more effective utilization of the medium, which is the outstanding technical problem in American broadcasting today.

Another important factor, from the standpoint of intensive development of the available frequency band, is frequency maintenance. In a system involving so many stations as are now operating in the United States, accurate maintenance of the assigned frequencies presents a very difficult problem. The maximum deviation permitted by the existing government regulations (± 500 cycles) is somewhat beyond the capabilities of the ordinary wavemeter and difficulty has been experienced in obtaining a satisfactory substitute. In the absence of adequate frequency control apparatus, very serious beatnote interference has been of frequent occurrence. During the past year, however, considerable improvement has been brought about by the extensive adoption of piezo-electric reference oscillators and automatic piezo-electric control. Equipment for the latter purpose capable of a relatively high standard of performance is now being offered commercially and it is probable that apparatus of this type will be installed in the near future by the majority of stations. Its use is expected to avoid entirely heterodyne interference on the "cleared" channels, where the beatnotes are those produced between the carriers of stations having adjoining frequency assignments. There is also reason to believe that the general adoption of such apparatus will materially improve conditions on the "shared" channels, each of which is occupied by several stations located at suitable distances, provided the assigned frequencies can be maintained with sufficient accuracy to preclude the reproduction of audible beats or other objectionable interference effects.

This problem of "synchronization," or preferably "common frequency operation," is beginning to receive considerable attention from all factors in the broadcasting industry. It promises important contributions in at least two directions:

(1) Improvements in the coverage of a common service area by two or more stations all broadcasting the *same* program;
(2) The attainment of minimum geographical spacings between stations operating on the same nominal frequency and broadcasting *different* programs.

The degree of frequency maintenance required for these two cases is apparently quite different. For case (1), the evidence indicates that very rigorous requirements must prevail. The most successful operations of this type have employed wire lines connecting the stations for the transmission of a base frequency from which the carriers were derived by means of harmonic generators. For case (2), however, there is reason to believe that comparatively wide limits will suffice.

Experience has shown that if the entertainment value of a program is not to be seriously impaired by interference, the ratio of wanted to unwanted carrier at the receiving point, in terms of field intensity, must be at least 100 : 1. From a relative interference standpoint, the significant factors are the wanted sidebands, the unwanted sidebands and the unwanted carrier, each of which produces a component in the detector output by interaction with the wanted carrier. With equal modulation at both stations, which is one of the conditions assumed, the ratio of the audio components due to the sidebands will, in general, be approximately the same as that between the carriers, or 100 : 1, representing a difference in level of 40 db. Due to the frequency difference between carriers, demodulation of one of the unwanted sidebands will result in the original signal with each of its elements shifted upward in pitch by an amount corresponding to this difference, while the other sideband will produce a signal which is similarly displaced in the reverse direction. The interfering signal may be badly garbled, therefore, but its disturbing effects insofar as enjoyment of the program is concerned will be substantially unaffected. The beatnote, which results from the interaction of the unwanted and wanted carriers, will be 6–10 db above this sideband interference level if average practice, as previously described, is followed. From this analysis, it appears that if the beatnote can be held to a value below the lowest frequency which it is desired to transmit and if one of the circuit elements of the reproducing system can be designed to provide some 10 db discrimination against the beat frequency, interference due to the latter can be so subordinated that the service areas of the stations involved will be defined by the limiting condition assumed for sideband interference, or a 100 : 1 ratio between carrier field intensities. Under these circumstances, no beatnote interference will be experienced in those areas where reasonably good service can be given. In adjoining regions, where the carrier ratio is less than 100 : 1, beatnote interference may continue to be observed but is of no importance since satisfactory reception in such areas is precluded by the sideband interference.

To meet the requirements outlined, it is probable that ultimately frequencies will have to be maintained to approximately 10 cycles, which would result in a maximum beatnote near the lower limit of aural frequency response. Such precision seems hardly necessary, however, under the conditions existing at the present moment. Almost all loud speakers now commercially available discriminate notably against frequencies below 100 cycles. A material improvement in beatnote conditions could probably be brought about, therefore, by the adoption of automatic control apparatus capable of maintaining the assigned frequencies to ± 50 cycles. Such performance is within the capabilities of the piezo-electric apparatus now available. Under the circumstances it is expected that considerable progress will be made in this direction during the coming year.

The foregoing considerations lead to the formulation of an important system requirement affecting receiving apparatus, which in this case includes both the radio receiver proper and the loud speaker. In a system involving a relatively large number of stations assigned to cleared and shared channels at 10-kc. intervals, such as exists in the United States, beatnote interference in the form of components at approximately zero cycles and at 10 kc. is an inherent characteristic. If a maximum frequency deviation of ± 10 cycles is accepted as the ultimate limit, in order to avoid such interference the receiving apparatus must be so designed that at frequencies below 20 cycles and above 9,980 cycles there will be introduced sufficient attenuation to suppress effectively the beatnotes likely to be encountered under any practical operating condition. Developed in this manner the proposition is more or less self-evident, but due to the rapidity with which the audio spectrum of broadcasting apparatus is being extended, some emphasis on the matter seems desirable.

Still another factor of importance from a system standpoint is control of radio harmonics. Spurious radiation of all types is inimical to intensive development and must be avoided. The harmonic problem presents unusual difficulties since efficiency requires that the tubes in the final power-amplifier stage be used in such a manner that relatively large harmonic voltages are impressed on the output circuit, yet the harmonic power radiated must be held to an extremely small amount. A measure of the purity of wave form required may be gained from the fact that a 5-kw. transmitter operating on a good antenna is capable of establishing an electromagnetic field of approximately 0.5 v per meter at a distance of one mile. Under the circumstances, a harmonic of 0.1 per cent represents a field intensity of 500 μv per meter at the same distance. Acceptable service in many areas is being obtained with field intensities of this order of magnitude. To bring the interfering field down to the static level would probably require reduction of harmonics to 0.01 per cent or less. From an apparatus standpoint, such performance represents a very difficult problem and it is questionable if it can be justified at the present time. Practice on this point is still in a state of flux, but there is reason to believe that some intermediate value, such as 0.05 per cent, will prove to be the proper solution, and will be applied to all broadcasting stations in the near future.

One circumstance that has undoubtedly contributed to the delay in formulating definite requirements concerning the control of harmonics has been the difficulty of obtaining suitable apparatus for the evaluation of such components in quantitative terms. Field strength measuring sets have recently been made commercially available, however, which are capable of covering the necessary range in frequency and intensity. A photograph of one of these sets is shown in Fig. 3. It consists essentially of a sensitive, stable superheterodyne receiver incorporating a calibrated attenuator at the input of the intermediate-frequency amplifier and a supplementary radio-frequency oscillator from which a voltage of the frequency of the station under measurement can be introduced in the antenna circuit. The operating characteristics of such an instrument have been described by Friis and Bruce.[3] By means of a series of removable loops and coils,

[3] H. T. Friis and E. Bruce, "A Radio Field-Strength Measuring System for Frequencies up to Forty Megacycles," *Proc. I. R. E.*, **14**, 507–519; August, 1926.

the set shown is capable of measuring field strengths ranging from approximately 0.01 to 7,000 mv per meter throughout the band 250 to 6,000 kc. Apparatus of this type is now in use by the radio inspection division of the Department of Commerce.

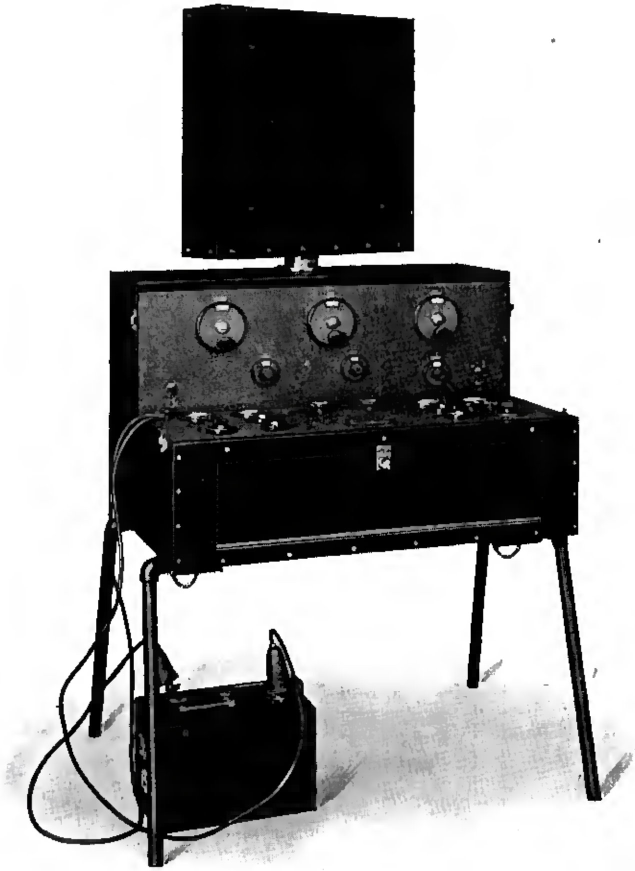


Fig. 3—Commercial field-strength measuring set. Range: 250–6,000 kc, 0.01–7,000 mv per meter.

In the light of this discussion of present trends in transmitter development, a brief description of some recent transmitting equipments may be of interest. A particularly noteworthy example of current practice is the 50-kw. Western Electric transmitter, one of which has been placed in service within the past few months by the Crosley Radio Corporation at Mason, Ohio. Views of this equipment are shown in Figs. 4, 5, 6, 7, and 8. The transmitter proper is shown in Fig. 4. As will be seen, it consists of seven panel units with a screen enclosure in the rear. The first unit on the left is the oscillator-modulator. This is essentially a low-power transmitter capable of an output of 50 watts and 100 per cent modulation. It is followed by three push-pull stages amplifying modulated radio-frequency power. The first power-amplifier stage, which employs two 250-watt tubes, occupies the second unit. The tubes for the second power stage, which are water cooled, and the associated tuned output circuit are contained in the third and fourth units, respectively. The final power stage, incorporating six water-cooled tubes each capable of a peak output of approximately 40 kw., occupies the fifth unit. The last two panels constitute the front of an electrically screened enclosure housing the output circuits for this latter stage. All of the panels are aluminum covered with several coats of black lacquer grained by rubbing with abrasive paper. A full complement of meters is provided, the cases of which are either grounded or mounted behind glass for the protection of the operating personnel. In designing the equipment, special consideration has been given to safety. Access to the apparatus in the rear of the panels can be had only through the door on the left which is held closed by a bolt operated by the handwheel shown. The rotation of this wheel opens the transmitter control circuits putting the equipment out of operation. It then grounds the high-voltage supply busses and finally withdraws the bolt. As an additional precaution a manually operated disconnect switch for the main power supply is provided just inside the gate which can be opened on entering. Access to some of the tubes is had by opening the glass windows in the various panels, but these are

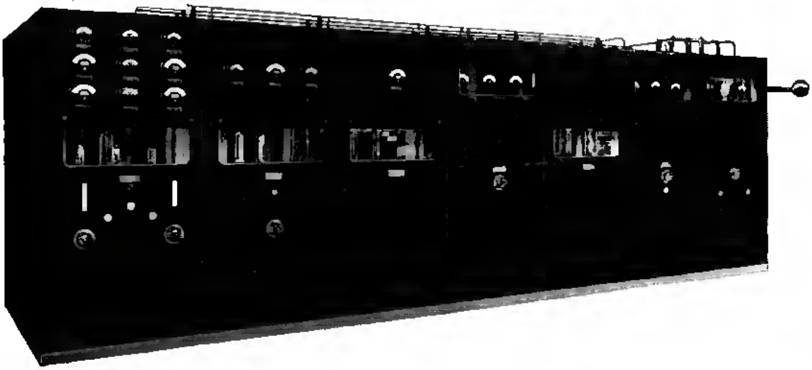


Fig. 4—Western Electric 7-A (50-kw) radio transmitter.
Oscillator-amplifier assembly.

secured by electrically operated latches unless the wheel is in the grounded position. Door switches are provided in the control circuits which prevent the transmitter from being placed in operation unless all doors and windows are closed.

The power panel and rectifier assembly is shown in Fig. 5. The general arrangement corresponds to that of the transmitter proper and similar safety features are provided. In the power panel, which is on the left, are centralized the necessary power distribution and control facilities. The equipment requires a 3-phase input of approximately 250 kw. at 440 volts. The control arrangement is such that the transmitter can be started and stopped by means of a single set of push buttons, the various circuits being energized in proper sequence by means of suitable relays and contactors. The central unit is a three-phase half-wave rectifier supplying power at 1,600 volts to the plates of the air-cooled tubes. The six-tube rectifier on the right supplies plate power at 17,000 volts for the water-cooled tubes. The filament and plate transformers and smoothing filter for the latter are located in the power room on the floor below. The filter consists of two units, one for each side of the push-pull circuit, employing a 6-μf condenser and a 12-henry inductance. Two 24-volt, 550-ampere direct generators (one a spare) supply power to the filament circuits.

Fig. 5—Power panel and rectifier assembly for 50-kw radio transmitter.

These machines are slot wound and employ composition brushes, a filter consisting of a 1-mh. inductance and four 1,000-μf electrolytic condensers being used to suppress commutator and slot ripples. Grid bias voltages are obtained from a 2-kw., 300-volt unit, which is also installed in duplicate. The only other rotating apparatus is that associated with the water-cooling system. The tubes are cooled by means of distilled water which is conducted to the anodes of the amplifier tubes through insulating hose coils. The total heat transferred by the cooling water is approximately 175 kw. A flow of 75 gallons per minute is maintained. Four 56-in. by 58-in. radiator units are employed, each consisting of a bank of copper tubes with spiral fins over which air is blown by a 37-in. fan. Ample radiator capacity is provided to maintain the water below 180 deg. F. under all atmospheric conditions.

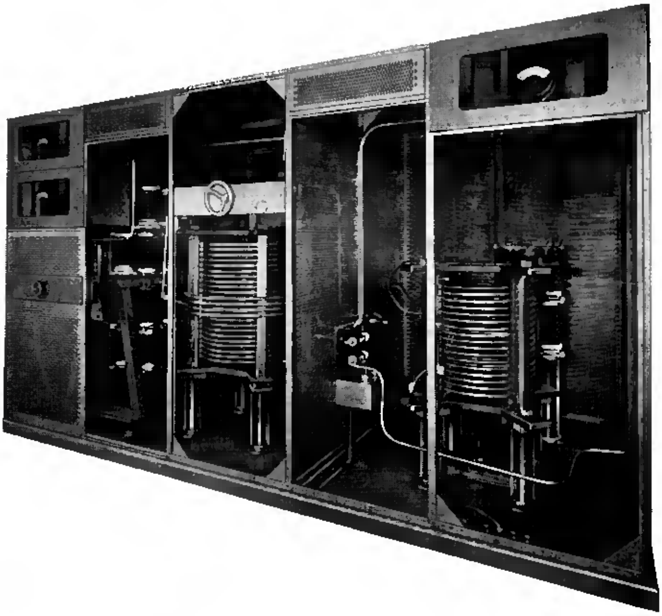


Fig. 6—Antenna coupling and tuning unit for 50-kw radio transmitter.

To promote antenna efficiency and to reduce the intensity of the electric field in the station building, the equipment is arranged to deliver its output to the antenna through a radio-frequency transmission line approximately 500 ft. long. The line is balanced to ground and is designed for a characteristic impedance of 600 ohms. The antenna coupling and tuning unit is shown in Fig. 6. It is intended for installation in a small building with a grounded copper roof located at the base of the antenna downlead. It consists of two tuned circuits, each housed in separate shielded compartments. In the photograph the doors and two of the screen panels have been removed to show the interior arrangement. The line is terminated by the parallel tuned circuit on the left which is inductively coupled to the antenna circuit to preserve an approximate balance to ground. The antenna is tuned by means of the series condenser and coil shown on the right. Accurate adjustment of the inductance of the coil is provided for by means of a short-circuited single-turn secondary which is located within the coil and arranged so that it can be rotated through approximately 90 deg. by the motor mounted on the floor beneath. The latter may be controlled from the operating room by a reversing switch placed on the right-hand panel of the transmitter assembly. A polyphase position indicator is provided to indicate the angle and movement of the secondary. The direct-current circuit of the thermal ammeter in the antenna circuit is also carried back to a bracket-mounted instrument on the end of the transmitter. These facilities permit the antenna tuning to be checked and adjustments

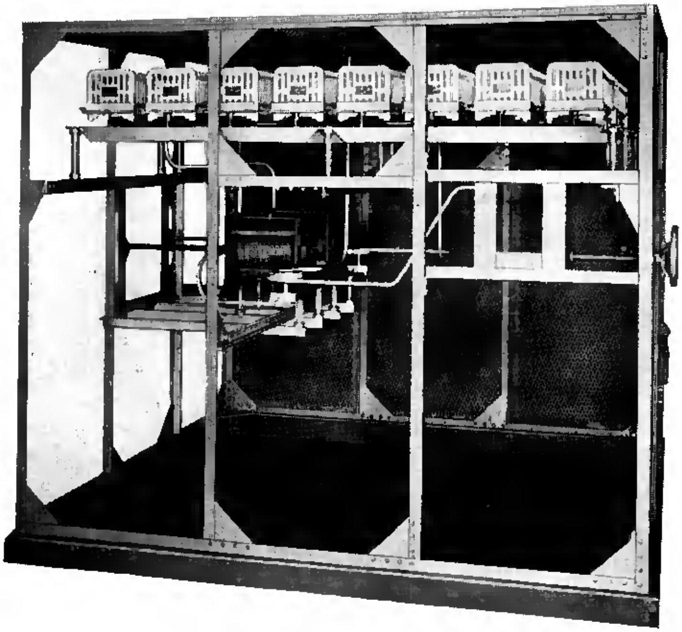


Fig. 7—Artificial antenna for 50-kw radio transmitter.

made to compensate for minor variations in antenna conditions without leaving the operating room.

Another feature of interest is the artificial antenna shown in Fig. 7. This unit is essentially a 600-ohm non-inductive resistance capable of dissipating approximately 75 kw. which can be connected to the output circuit of the final power amplifier stage in place of the transmission line. The heat dissipating elements consist of a series of woven wire grids mounted in the units at the top of the framework. The resistance of these grids is substantially independent of frequency, but the combination presents a slight inductive reactance which is compensated for by means of the condenser and coil combination shown. These elements are inserted into the circuit symmetrically in order to maintain an approximate balance to ground. The structure is completely shielded and is fitted with safety door and grounding switches similar to those already described.

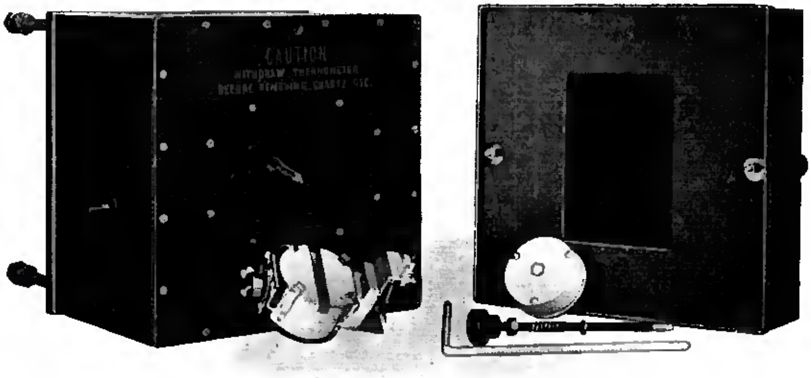


Fig. 8—Piezo-electric crystal mounting and temperature control apparatus.

The piezo-electric crystal mounting and temperature-control apparatus which is a part of the oscillator-modulator unit is shown in Fig. 8, dismantled to facilitate inspection. The quartz plates employed are approximately one and a quarter inches square and are cut parallel to one of the faces of the natural rock crystal. This plate is mounted between two lapped metal plates and covered with a porcelain cap carrying a terminal to which the upper electrode is connected by means of a short section of metal foil. The mounted crystal is supported by a brass block, through the center of which extends a spiral bimetallic thermostat. The top of the block is also lapped and the crystal mounting is secured to it by means of the four springs shown. The heating element consists of a winding of resistance wire inserted in the block concentric with the thermostat. The assembly is mounted in a thermally insulated box, shown on its side in the photograph. Two of these units are provided, one located on each side of the oscillator-modulator unit directly below the window. A detachable handle for adjusting the contacts of the thermostat and a suitable thermometer extend through the box to the front of the panel. The brass mounting block is provided with a groove to receive the bulb of the thermometer. The thermostat does not operate directly in the heater circuit but controls the grid bias of a vacuum tube in the plate circuit of which a suitable relay is placed. The quartz plates are ground to oscillate at the assigned frequency at approximately 50 deg. C, and the final adjustment is made by varying the operating temperature. The temperature coefficient of the plates varies from 30 to 100 parts in a million per deg. C. The degree of constancy attained necessarily depends on the diligence of the operating personnel. With proper maintenance the maximum deviation has been held to ± 30 cycles for long periods of time.

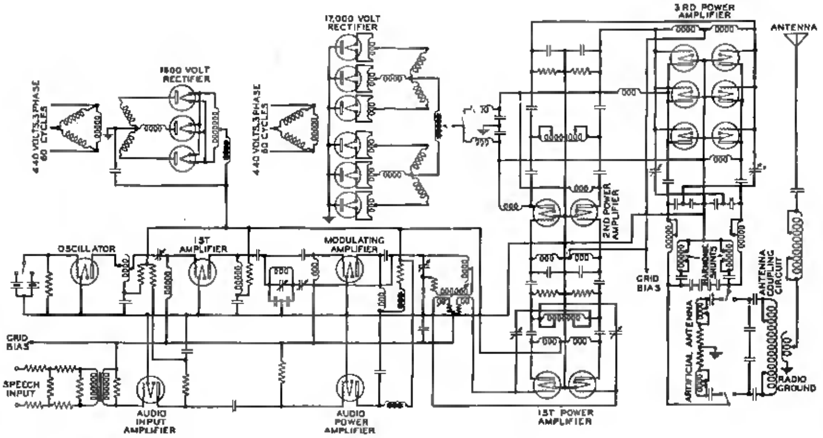


Fig. 9—Simplified circuit schematic of 7-A (50-kw.) radio transmitter.

A simplified circuit schematic is shown in Fig. 9. Features of the electrical design are the modulation system, the push-pull amplifier stages with cross neutralization, the capacity coupling arrangement used to facilitate control of parasitic oscillations, and the provisions for the suppression of harmonics. The modulating amplifier is a 50-watt tube operating at 750 volts. The audio power stage employs a 250-watt tube at 1,500 volts. In this manner, ample audio-frequency voltage and power are provided to effect complete modulation without distortion in the audio tube. With so powerful an equipment, the suppression of radio-frequency harmonics to a satisfactory degree becomes a difficult problem. The push-pull circuits, capacity coupling, three tuned circuits in cascade, shielding of all coils, and the two tuned shunts adjusted to the second harmonic which are connected between each side of the transmission line and ground all contribute to superior performance in this respect. The amplitude of the harmonics radiated, as determined by field strength measurements, is less than 0.03 per cent.

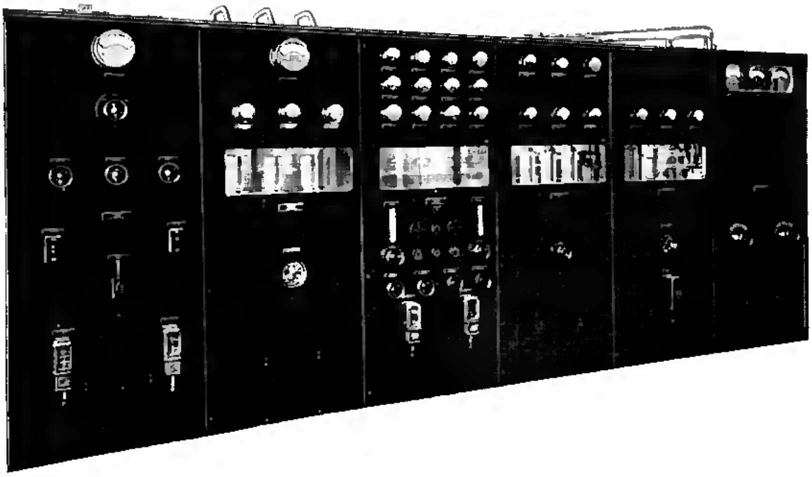


Fig. 10—Panel assembly for Western Electric 5-C (5-kw.) radio transmitter.

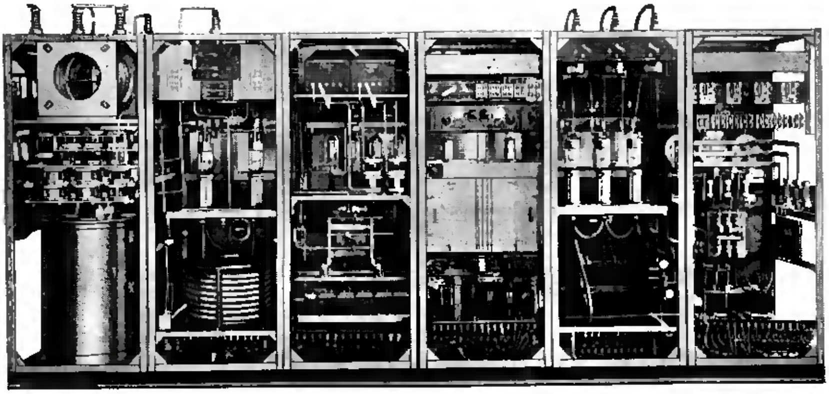


Fig. 11—Rear view of panels in 5-C radio transmitter.

A 5-kw. equipment of similar general design is shown in Figs. 10

and 11. It consists of six units: A power panel, a 10,000-volt rectifier for the water-cooled tubes, a piezo-electric oscillator unit, an intermediate amplifier unit, a power amplifier unit employing two 10-kw. tubes, and an output unit. An air-cooled transformer for the rectifier, the associated filter, and an artificial antenna are assembled in a screened enclosure in the rear of the panels. Three motor-generator sets are provided to supply filament power, grid bias, and plate power for the air-cooled tubes. A 3-phase power input of 30 kw. at 220 volts is required. The equipment is capable of fidelity in transmission comparable with that of the 50-kw. unit. The amplitude of the harmonics radiated is held to approximately 0.2 per cent.

Fig. 12—Western Electric 6-B (1-kw.) radio transmitter.

A 1-kw. equipment of the same type is shown in Figs. 12 and 13. It involves only two panels, a piezo-electric oscillator unit and an amplifier unit. The final power stage employs a 4-kw. water-cooled tube. Two motor generators are used, one supplying 24 volts and

250 volts for filaments and grid bias, the other 2,000 volts and 4,000 volts for the plates of the air-cooled and water-cooled tubes, respectively. A power input of 10 kw. is required.

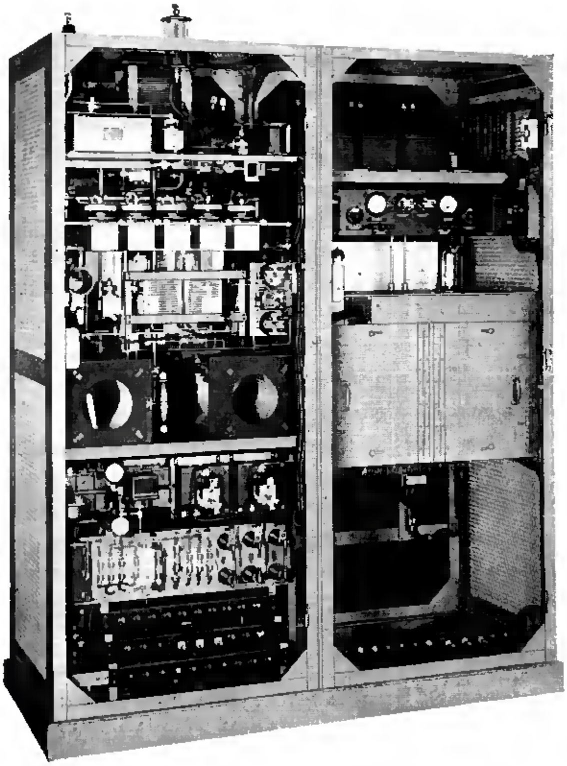


Fig. 13—Rear view of 6-B radio transmitter.

RADIO TRANSMISSION PHENOMENA

Radio transmission phenomena in the broadcasting band have been given considerable study, and the general nature of the effects likely to be encountered are fairly well understood. Important contributions have been made by Bown and Gillett, by Bown, Martin, and Potter, by Goldsmith, and by Espenschied.[4] The second paper referred to is particularly noteworthy on account of the insight which it affords into the complexities of the process of transmission and the evidence which it presents concerning the injurious effects of frequency modulation. The latter has not yet fully received the attention which it deserves; many otherwise well designed transmitters

[4] See attached list of references.

are still in operation that are subject to frequency changes of the order of ± 1,000 cycles during modulation. This condition is not only conducive to impaired fidelity at moderately distant receiving points, but it increases interference and precludes successful common frequency operation. Fortunately, the use of automatic frequency control apparatus in its present form is effective in minimizing this effect as well as in limiting frequency variations of much longer period. It is probable, therefore, that with the more general use of automatic piezo-electric control, this matter will rapidly cease to be a problem.

As might be expected, the attention being given to intensive development has materially stimulated interest in transmission. There is a very evident need for much information of a more quantitative nature than is now available. Data concerning attenuation over city and rural areas as a function of frequency, suitable separations between stations of various powers operating on a common carrier frequency, allowable distances between transmitting stations and nearby populous communities, relative day and night ranges, relative summer and winter ranges, time of the day and season of the year at which the transition occurs, and other questions of a similar nature have become of great practical importance. The problem is rendered particularly difficult by the range in climatic, topographic, and cultural conditions which exist in the United States. Under the circumstances, there are excellent opportunities for important work in this field.

A significant tendency disclosed by recent measurement work in a number of city areas is public acceptance of and demand for field intensities which a few years ago would have been considered objectionably high. For some time it has been more or less generally agreed that a field intensity of 10 mv. per meter would afford a satisfactory high-grade broadcasting service. Recently, however, in spite of increased effectiveness due to higher degrees of modulation and in spite of continued improvement in the sensitivity of commercial receiving sets, stations establishing field strengths of 10–15 mv. per meter have been greatly handicapped in competing with others capable of producing 30–50 mv. per meter in the same areas. In several densely populated districts measurements have disclosed field intensities of 300–500 mv. per meter without any noteworthy number of complaints provided the programs were of a high character. There is little to indicate whether this tendency is the result of a decreased interest in distant stations, a desire for higher standards in reproduction involving lower noise levels, or a combination of these factors with others, but it is evidently a matter which must be given careful consideration in engineering future installations.

It is interesting to contrast this situation with that existing in some of the large rural districts as exemplified by the recent survey of conditions in the Middle West by Jansky.[5] Here over large areas acceptable service is being obtained with field strengths of 50 and 100 μv per meter. Giving due consideration to the difference in noise levels, which is undoubtedly a factor of great significance, such a discrepancy can only be reconciled on the basis of a vast difference in service standards. That such conditions will be allowed to continue for any considerable period of time is very doubtful. This is further evidence indicating that the movement toward more powerful stations is technically sound.

One phase of the transmission problem which deserves increased attention is antenna performance and design. It is an interesting circumstance that while the accurate rating of broadcasting stations is a matter of great practical concern to the industry, to date consideration has been confined to the power delivered to the antenna. Variations in the efficiency of the latter have been almost entirely neglected in spite of the fact that, due to this cause, the power actually radiated can be shown to vary through a range of four to one, or greater. There is little doubt that stations should be rated, either directly or indirectly, in terms of field intensity measurements. That such a system of rating has not already been put into effect is probably due to the lack of suitable measuring apparatus. With such equipment now available, rapid progress in this direction is expected.

An interesting feature of current American practice with respect to broadcasting antennas is a definite tendency toward the use of higher supporting structures. For the past few years, most of the towers erected have been from 150 to 225 ft. in height. Several of the more recent stations are employing 300-ft. towers, and it is not improbable that some 400-ft. structures will be put up in the near future. Since the natural frequency of grounded steel towers of these dimensions falls in the broadcasting band and may approximate the assigned operating frequency, low-capacity porcelain insulators are inserted at the base. The latter effect a considerable increase in the natural frequency of the towers and preclude serious distortion in the field intensity pattern due to heavy induced currents in the steel. The antennas themselves are of such dimensions that the current antinode is positioned well up on the vertical section. The effect is to concentrate the radiated power along the ground plane and to increase materially the field intensity in the local service area. Such antenna systems promise a better economic balance between the in-

---

[5] See attached list of references.

vestment for generating modulated radio-frequency power and that for radiating it.

REFERENCES

RALPH BOWN, CARL R. ENGLUND, AND H. T. FRIIS. Radio Transmission Measurements. Proc. I. R. E., **11,** 115; April, 1923.

D. G. LITTLE. KDKA Telephone Broadcasting Station of the Westinghouse Electric and Manufacturing Co., East Pittsburgh, Penna. Proc. I. R. E., **12,** 255; June, 1924.

RALPH BOWN AND G. D. GILLETT. Distribution of Radio Waves from Broadcasting Stations over City Districts. Proc. I. R. E., **12,** 395; August, 1924.

EDWARD L. NELSON. Transmitting Equipment for Radio Telephone Broadcasting. Proc. I. R. E., **12,** 553; October, 1924.

JULIUS WEINBERGER. Broadcast Transmitting Stations of the Radio Corporation of America. Proc. I. R. E., **12,** 745; December, 1924.

RALPH BOWN, DELOSS K. MARTIN, AND RALPH K. POTTER. Some Studies in Radio Broadcast Transmission. Proc. I. R. E., **14,** 57; February, 1926.

ALFRED N. GOLDSMITH. Reduction of Interference in Broadcast Reception. Proc. I. R. E., **14,** 575; October, 1926.

LLOYD ESPENSCHIED. Radio Broadcast Coverage in City Areas. *Bell Syst. Tech. Jour.*, VI, 117; January, 1927.

D. K. MARTIN, G. D. GILLETT, AND I. S. BEMIS. Some Possibilities and Limitations in Common Frequency Broadcasting. Proc. I. R. E., **15,** 213; March, 1927.

KNOX MCILWAIN AND W. S. THOMPSON. A Radio Field Strength Survey of Philadelphia. Proc. I. R. E., **16,** 181; February, 1928.

I. F. BYRNES. Recent Developments of Low Power and Broadcasting Transmitters. Proc. I. R. E., **16,** 614; May, 1928.

P. P. ECKERSLEY. The Design and Distribution of Wireless Broadcasting Stations for a National Service. Proc. Wireless Section, I. E. E., **3,** 108; June, 1928.

H. M. O'NEILL. Characteristics of Certain Broadcasting Antennas at the South Schenectady Development Station. Proc. I. R. E., **16,** 872; July, 1928.

S. W. EDWARDS AND J. E. BROWN. The Use of Radio Field Intensities as a Means of Rating the Outputs of Radio Transmitters. Proc. I. R. E., **16,** 1173; September, 1928.

C. M. JANSKY, JR. Some Studies of Radio Broadcast Coverage in the Middle West. Proc. I. R. E., **16,** 1356; October, 1928.

Wire Line Systems for National Broadcasting [1]

By A. B. CLARK

The interconnecting of radio broadcasting stations by special telephone lines for the simultaneous broadcasting of radio programs began on a commercial basis in 1923. Today well over 30,000 miles of program transmission circuits are in use in the United States and transcontinental broadcasts by means of such wire lines are a daily occurrence.

The paper first states the radio limitations which make wire lines necessary for broadcast coverage of large nations. A map and data are given showing the present broadcasting chains in the United States and indicating the extent of their use. An explanation is given of why program transmission circuits must have transmission characteristics materially different from message telephone circuits and a brief discussion of some of the important transmission characteristics of such circuits, including particularly "frequency range" and "volume range." The present chains in the United States which are made up almost entirely of open-wire circuits on a voice-frequency basis are briefly described. The manner in which these chains are tested and the way control is exercised are also indicated. To exercise this control requires an elaborate network of telegraph wires now aggregating over 40,000 miles and a corps of special men over 300 in number.

WHAT we are here considering, as an important factor in promoting national solidarity, is the tying together of a whole nation so that a single broadcast will instantly reach even the most remote points. Radio broadcasting stations (employing the more generally used frequencies) are essentially local distribution centers serving effectively points up to 50 miles (80 kilometers) or, in favorable cases, 100 miles (160 kilometers) or more from the radio transmitter. For the larger nations it is evidently necessary to make division into areas, locating a radio transmitter in each area for its coverage, and then to provide a network of circuits connecting the transmitters in the various areas with the point at which the broadcast originates. At the present time wire telephone systems are employed almost exclusively for this national distribution of broadcasts. It is the purpose of this paper to discuss the wire networks which are now being provided in the United States by the Bell Telephone System.

In the United States at the present time (January 15, 1929) programs are being regularly distributed over extensive wire networks or "chains" as indicated on the map of Fig. 1.[2] The various chains

---

[1] Presented before the World Engineering Congress at Tokio, Japan, October, 1929, Proc. of the I. R. E., November, 1929.

[2] This map has been revised to show the network chains as of September 1, 1929.

are usually referred to by colors and are so designated on the map. As a regular procedure most of these chains operate about six hours each day. Following are the numbers of radio stations served by each chain together with the lengths of telephone circuit involved. (An additional chain which operates only one hour each week is not included.)

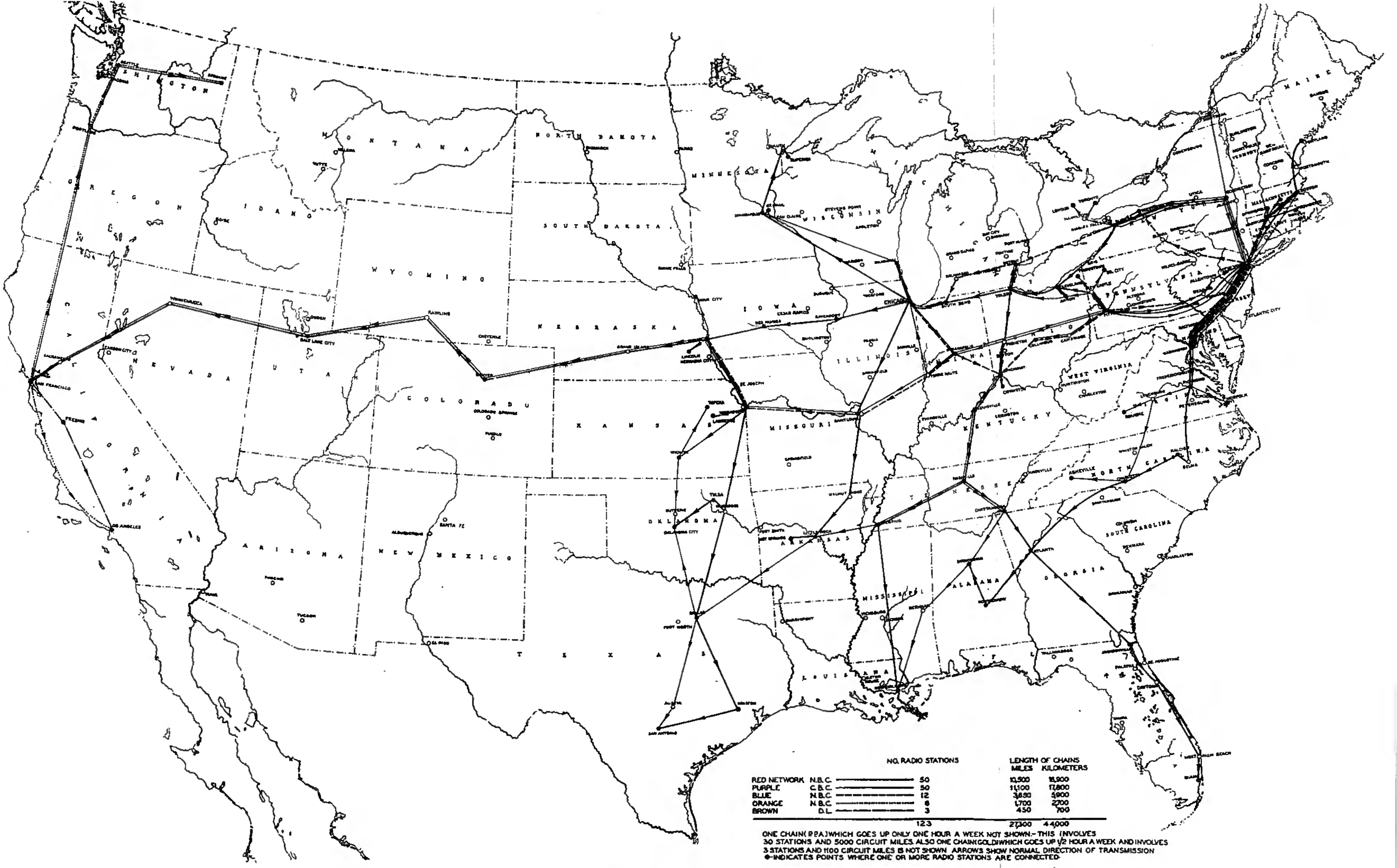
	Radio Stations	Telephone Circuit Miles		
Red network[3]	41	10,500	16,600	kilometers
Purple network	41	8,450	13,600	"
Blue network	12	3,650	5,900	"
Green network	8	3,600	5,800	"
Orange network	5	1,700	2,700	"
Brown network	3	450	700	"
Total	110	28,150	45,300	"

[3] See table on Fig. 1 for revised data as of September 1.

On occasions when events of particular importance take place, several of the regular chains may be merged together and additional circuits added so as to pick up programs from various parts of the country. For example, on November 5, 1928, the evening before the United States presidential election, the networks shown in Fig. 2 were in operation, about 85 radio stations being included. At various times during this evening, five separate programs were broadcast from several different points in New York City; Palo Alto, California; Little Rock, Arkansas; and Pittsburgh, Pennsylvania. The United States was thus virtually one great auditorium, with listeners estimated as no less than fifty million.

From the technical standpoint, program transmission circuits are, of course, very different from message telephone circuits. In the first place, message telephone circuits must be arranged so that to and fro conversations can take place practically instantaneously. Program transmission circuits on the contrary are single-direction transmission circuits. They are, therefore, not complicated by problems of electrical echo, singing and the like, which are ever present with long message telephone circuits. However, although free from the problems of two-way working, the design and operation problems of program transmission circuits are by no means easy as compared with those of message telephone circuits. On the contrary, in many respects, these problems are considerably more difficult, the reason being that the requirement as to approach to absolute fidelity of reproduction is much more severe than for message telephone circuits.

A frequency band width of 2,500 cycles furnishes, if properly utilized, a telephone circuit over which speech is transmitted very clearly so that conversations may be easily carried on. This band is not



		NO. RADIO STATIONS	LENGTH OF CHAINS MILES	KILOMETERS
RED NETWORK	N.B.C.	50	10,500	16,900
PURPLE	C.B.C.	50	11,100	17,800
BLUE	N.B.C.	12	3,650	5,900
ORANGE	N.B.C.	8	1,700	2,700
BROWN	D.L.	3	450	700
		123	27,300	44,000

ONE CHAIN (P.P.A.) WHICH GOES UP ONLY ONE HOUR A WEEK NOT SHOWN.– THIS INVOLVES 30 STATIONS AND 5000 CIRCUIT MILES. ALSO ONE CHAIN (GOLD) WHICH GOES UP 1/2 HOUR A WEEK AND INVOLVES 3 STATIONS AND 1100 CIRCUIT MILES IS NOT SHOWN. ARROWS SHOW NORMAL DIRECTION OF TRANSMISSION
●-INDICATES POINTS WHERE ONE OR MORE RADIO STATIONS ARE CONNECTED.

Fig. 1—Routes of Bell System Program Network Chains in the United States as of September 1, 1929.

the amount of distortion which theory and experience indicate should be expected. Then, final adjustments are made by certain specially provided adjustable parts in accordance with the overall measurements. Such overall tests and adjustments are, in general, made daily.

In setting up these circuits, another important consideration is that each amplifier carry its proper load or, in telephone parlance, each amplifier deliver to its associated line the proper output level. To insure this, diagrams are prepared in advance, showing the desired transmission levels at each repeater, a typical diagram being shown in Fig. 3. In setting up the circuits, the repeater gains are first set to values which theory and experience indicate should result in conditions as shown in the prescribed transmission level diagram. Testing current is then applied to the sending end of the circuit and sensitive measuring devices are applied at the output of each repeater. If the results of these measurements do not accord with the transmission level diagram, suitable adjustments are then made.

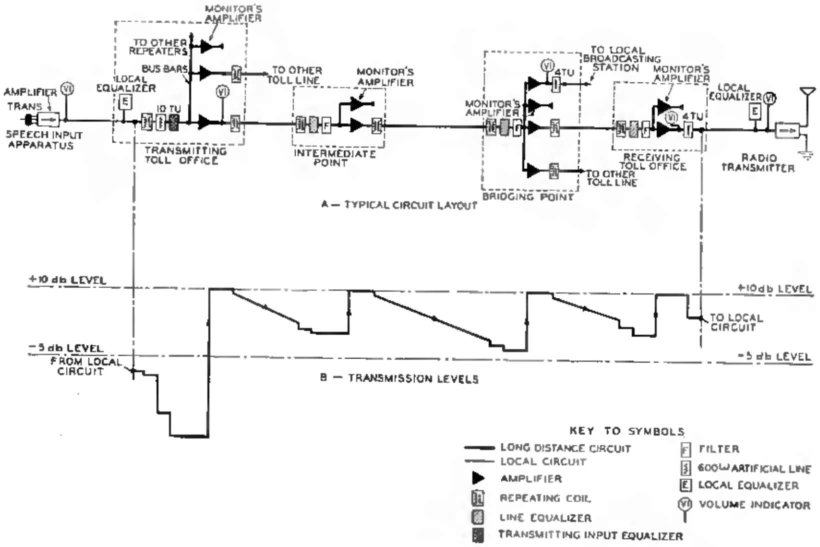


Fig. 3—Typical Circuit Layout and Transmission Level Diagram of Program Network Circuits.

In building up the large chains which tie together a considerable number of radio transmitters, wire distributing centers are provided at strategic points. Figure 4 shows the circuit layout of the various chains which have been referred to and indicates in a general way how produce an undue amount of disturbance in neighboring circuits which may be transmitting other programs or telephone messages. The designer is also concerned lest when the program power is weak the programs be unduly interfered with by noise or crosstalk from other circuits. He must particularly consider the noise and crosstalk which may be heard during pauses in programs. During such pauses it is very annoying to the listeners to hear a background of noises of various sorts and it is essential that the listeners be unable during such pauses to pick up intelligible speech from telephone message circuits crosstalking into the program circuit.

At the present time generally satisfactory results are being obtained in transmitting the volume range of about 30 decibels (3.4 nepers). Considerably more must be done both in the radio and in the wire systems, however, before there can be transmitted volume ranges comparable with those put out by symphony orchestras, high-grade artists, and the like.

Having indicated in a general way the requirements of program transmission circuits, there will next be described the wire systems which are now in use in the United States.

The present-day program transmission circuits in the United States are "on a voice-frequency basis," which means that the waves transmitted over the circuits are essentially copies of the sound waves impinging on the microphones. Most of the circuits now being provided are carried by the familiar open wires, usually copper wires 0.165 inch (4 mm.) in diameter spaced about 1 foot (30 cm.) apart on the crossarms. The transmission properties of an open-wire pair without loading are well suited for program transmission purposes since the distortion is comparatively small although it is far from negligible. Spaced at intervals on these circuits, averaging roughly 150 miles (240 kilometers) apart, are one-way repeaters or amplifying devices. Along with these amplifiers are other electrical devices for counteracting the distortion introduced by the open-wire circuits, incidental cables involved, etc. Other one-way repeaters are provided at the terminals of the circuit. Considerable technical refinement is, of course, involved in the design of these amplifiers and of the auxiliary apparatus associated therewith which cannot be gone into here.

In setting up the program transmission circuits, an important part of the work consists in making measurements at different single frequencies within the band which it is desired to transmit over the circuit. Before making such overall measurements, the amplifiers and auxiliary apparatus are so adjusted locally as to compensate for adequate, however, for program transmission because of the different character of the transmitted material. The bulk of present-day broadcast programs consists of musical selections, including a fair amount of high-grade material. To reproduce music, and particularly high-grade music, in a pleasing manner calls for a materially widened band. This wider band also gives a high degree of naturalness to speech which is particularly desirable when loudspeakers are used for reception.

At the present time in the United States the frequency band which is transmitted over the long distance program chains extends from about 100 cycles to about 5,000 cycles. It is, of course, possible to transmit an even wider band than this, although the cost of the circuits will, of course, increase as the band is widened. In considering how wide the band should be, the complete system, including pickup apparatus, wire transmission line, radio transmitters, radio transmission paths through the ether, radio receiving apparatus and loud speakers must be considered. It seems probable that as the art progresses a band wider than the above will be found desirable. On the wire line systems, development work is going forward looking toward the possibility that such wider bands may be found desirable in the future. At the lower frequencies, where most people consider that improvement is particularly desirable, consideration is being given to the possible extension of the band down to 50 cycles and possibly lower. Consideration is also being given to the possible addition of two or three thousand cycles to the top of the band.

In addition to this broad band transmission requirement, program transmission circuits must be designed to handle wide ranges of volume, particularly for the transmission of musical programs. Much of the enjoyment in listening to good music appears to come from the ranges of volume, so that in order to deliver such musical programs properly these ranges of volume must be preserved in large part at least. At the present time the volume ranges are "compressed" somewhat by adjustment of amplification under control of an operator at the pickup point. This tends to make easier the radio transmission problem as well as the wire transmission problem. The range of volume which is now delivered, as read by a "volume indicator" (a meter which roughly indicates the peaks), is of the order of 30 decibels (3.4 nepers), which means that during the fortissimo parts of programs the power which is transmitted is about 1,000 times as great as it is during the pianissimo portions.

The designer of the wire circuits must be concerned lest during those periods when the program power is strong, the program circuits

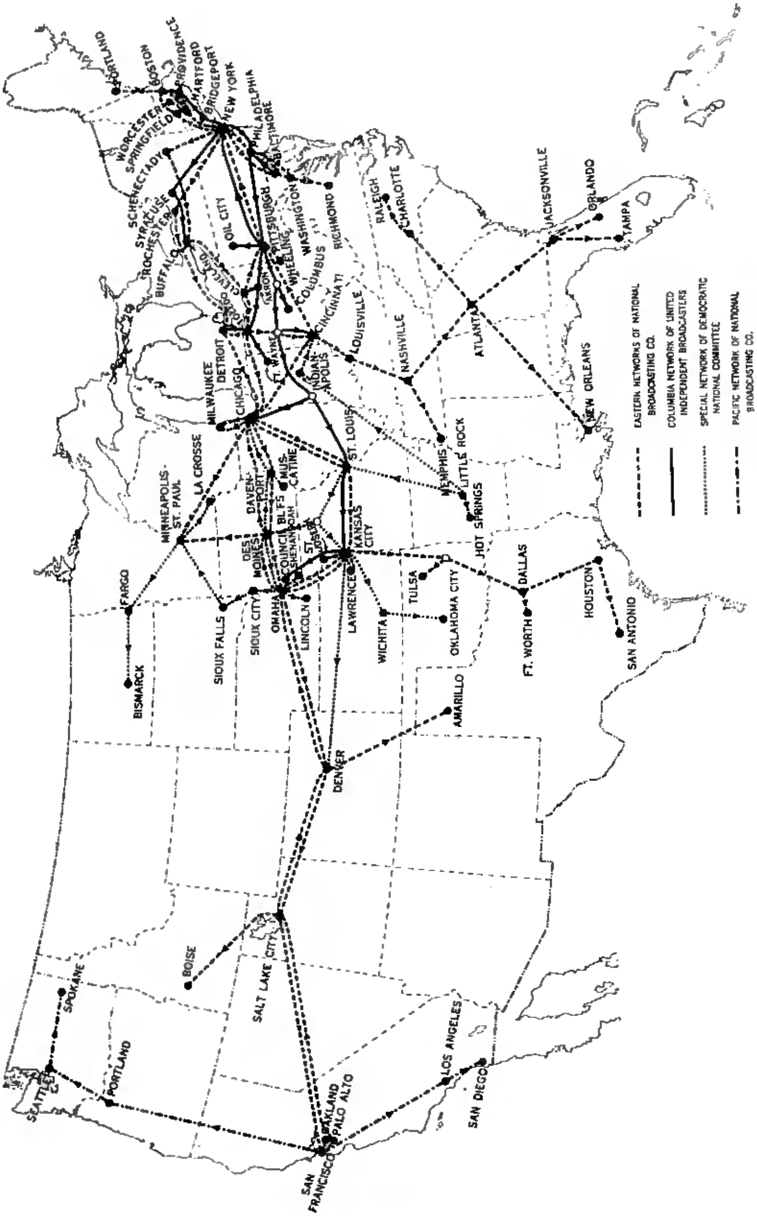
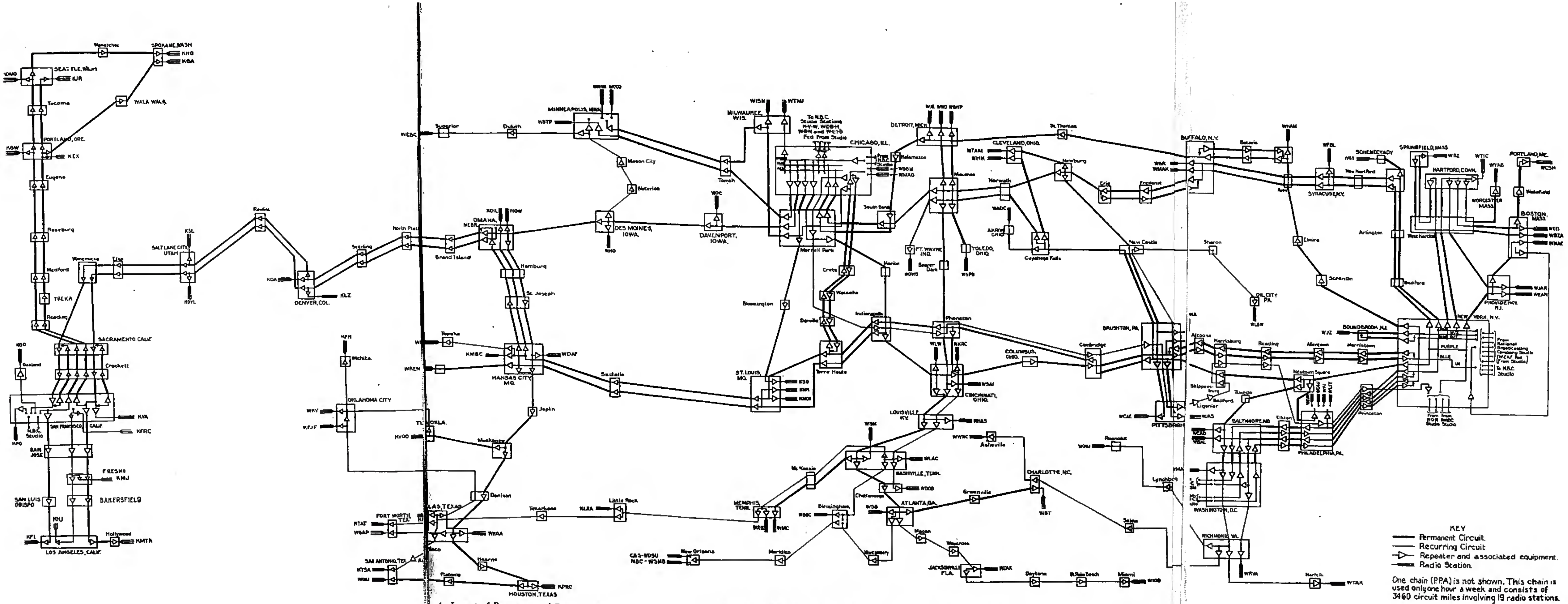


Fig. 2—Bell System Program Networks Operating on Last Night before 1928 Presidential Election.



One chain (PPA) is not shown. This chain is used only one hour a week and consists of 3460 circuit miles involving 19 radio stations.

4—Layout of Permanent and Recurring Program Network Circuits with Associated Repeater Equipment Shown on the Basis of Normal Direction of Transmission as of January 15, 1929.

the various chains are interconnected and arranged for switching at certain distributing centers.

In the United States the largest distributing center is, naturally, in New York City, since the bulk of the program material originates at that point. At such a distributing center a special collection of various forms of equipment is provided consisting of one-way amplifiers, loud speakers, multifrequency oscillators, various forms of transmission measuring devices and miscellaneous apparatus. The photograph of Fig. 5 shows a portion of the program layout in the New York long distance telephone office as of January 15, 1929. The various bays at the left carry the line apparatus associated with branches of various chains. In the rear are located the transmission measuring apparatus and multifrequency oscillators. In the foreground are the terminals of various telegraph order wires.

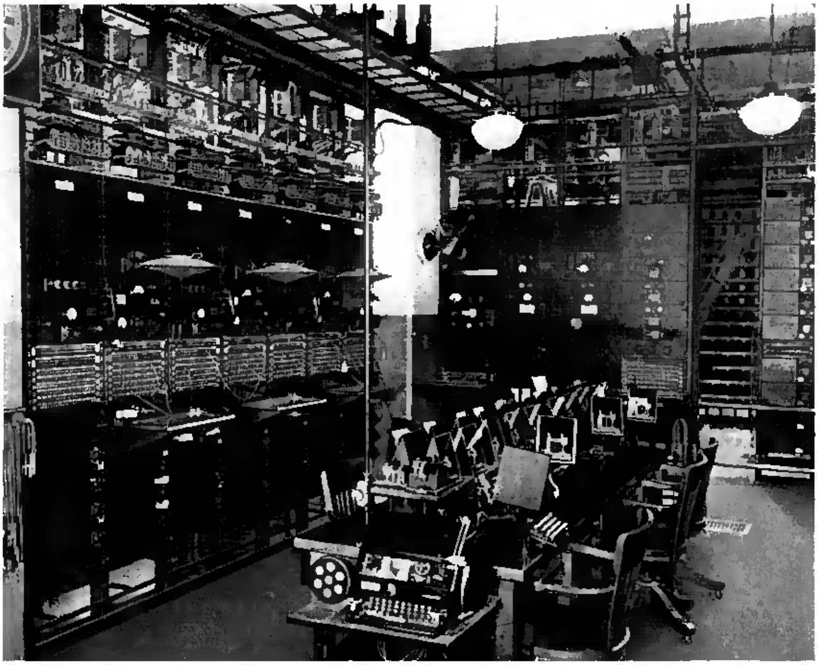


Fig. 5—Portion of Program Apparatus Layout in New York Long Distance Telephone Office as of January 15, 1929.

In transmitting programs over a wire network, as has been pointed out above, it is important that the volume range be held within proper limits. It is one of the obligations of the one who "picks up" the program to hold his range of volume between proper limits. At the central distributing point those in charge of the wire circuits usually find it desirable to make checks from time to time to insure that the proper range of volume is maintained. This checkup is made by means of a device known as a "volume indicator" similar to the one which the program supplier uses for purposes of regulating his volume range. Other volume indicators are provided at various strategic points in the wire network in order to insure that the proper range of volume is reaching these points. In addition to regularly making these observations by means of volume indicators, loudspeaker monitoring observations are continually made at practically all repeater points.

The results of these observations are transmitted back to the control points periodically by means of telegraph order wires so that the control operator knows at all times the condition of transmission at every point in his territory.

With the network chains grown to such vast proportions as indicated in Figs. 1 and 4, it is essential that the system for controlling the networks be such that all points involved be in instant communication with certain designated control points. To accomplish this, the United States has been divided into four areas, each area of which is under the control of a distributing center or control station. The four control stations in the United States at present (January 15, 1929) are, New York covering the eastern section, Chicago the western section, Cincinnati the southern section, and San Francisco the Pacific Coast section. Each of these control points is connected to every repeater point in its area by means of telegraph order wires and in addition is connected to every radio station in the area served by the networks under its control. The various control points are also connected together by means of order wires and arrangements are provided so that New York can be placed in communication with any of the radio stations in the United States which are served by the chains. The total telegraph wire mileage employed for this service is now approximately 43,000 miles (70,000 kilometers).

A large corps of specially trained telephone men is needed to properly supervise the transmission performance of the chains as well as to take care of the switching and general coordination work involved. At present, about 300 men are employed in the United States for this service, these men, of course, being in addition to those who care for the regular wire and equipment maintenance.

ACKNOWLEDGMENT

Acknowledgment is made to Mr. H. S. Hamilton for considerable assistance in connection with the preparation of the text and particularly of the drawings, and to Mr. G. S. Bibbins and Mr. H. C. Read for furnishing most of the statistical data.

Notes on the Heaviside Operational Calculus

By JOHN R. CARSON

This paper briefly discusses the following topics: (1) the asymptotic solution of operational equations; (2) Bromwich's formulation of the Heaviside problem, and its relation to the classical Fourier integral; and (3) the existence of solutions of the operational equation. The paper closes with some general remarks on the interpretation of the operator and the operational equation, emphasizing the purely symbolic character of the latter.

THE large amount of work done in the past thirteen years, starting with important papers by Bromwich[1] and K. W. Wagner,[2] has served to remove whatever mystery may have surrounded the Heaviside operator, and has placed his operational calculus on a quite secure and logical foundation. However, certain phases of the problem still do not appear to the writer to have as clear or adequate treatment as perhaps might be desired; these it is the object of the present paper to discuss. The topics dealt with are (1) the asymptotic solution of operational equations; (2) Bromwich's very important formula and its relation to the classical Fourier integral; and (3) the existence of solutions of the operational equation.

In the following it will be assumed that the reader has a general acquaintance with the Heaviside operational calculus as well as the Fourier integral, but a brief sketch of the former may not be out of place. It will be recalled that the Heaviside processes were originally developed in connection with the solution of electrical problems:[3] more precisely, the determination of the oscillations of a linearly connected system specified by a set of linear differential equations with constant coefficients or a partial differential equation of the type of the wave equation. This system is supposed to be in a state of equilibrium at reference time $t = 0$, when it is suddenly acted upon by a 'unit' force (zero before, unity after time $t = 0$); the subsequent behavior of the system is required. In the solution of this problem, Heaviside's first step was the purely formal and symbolic one of replacing the differential operator $\partial/\partial t$ by the symbol p, thereby

---

[1] "Normal Coordinates in Dynamical Systems," *Proc. Lond. Math. Soc.* (2), 15, 1916.

[2] "Über eine Formel von Heaviside zur Berechnung von Einschaltvorgänge," *Archiv. Elektrotechnik*, Vol. 4, 1916.

[3] Since this paper is addressed largely to physicists and engineers, we shall employ to some extent the language of circuit theory rather than pure mathematics; no loss of essential generality is involved.

reducing the differential equations to an algebraic form, the formal solution of which we shall write

$$h = \frac{1}{H(p)}, \qquad t \geqq 0. \tag{1}$$

Here $h = h(t)$ is the variable with whose determination we are concerned and $H(p)$ is the Heaviside function, derived as stated from the differential equations of the problem. This equation is as yet purely symbolic, and its conversion into an explicit solution for h, as a function of t, constitutes the Heaviside problem.

Bromwich[1] formulates the problem as the infinite integral

$$h(t) = \frac{1}{2\pi i}\int_{c-i\infty}^{c+i\infty} \frac{e^{pt}}{pH(p)}\,dp. \tag{2}$$

The writer's formulation of the problem is, that h is uniquely determined by the integral equation[4]

$$\int_0^\infty h(t)e^{-pt}dt = \frac{1}{pH(p)} \tag{3}$$

This equation is valid for all values of p, for which its *real part* is greater than some finite constant c; c must be at least large enough to make the infinite integral converge. In the majority of physical problems this constant may be taken as 0; in some, however, the equation is valid only when c is greater than some finite constant.

The equivalence of (2) and (3) is very easily established in a number of ways; perhaps the simplest is to show, following March,[5] that (2) is the formal solution of (3). Either can be deduced from the other. The Bromwich solution can, of course, be derived directly from the Heaviside problem, as shown below.

I

One of the most interesting and perhaps the least generally understood of Heaviside's methods of solving the operational equation is the process whereby he derives a series solution, usually divergent and asymptotic, in inverse fractional powers of t. What I have termed the Heaviside Rule[6] for deriving this type of solution may be formulated as follows:

---

[4] "The Heaviside Operational Calculus," *B. S. T. J.*, 1922; *Bulletin Amer. Math. Soc.*, 1926.

[5] "The Heaviside Operational Calculus," *Bulletin Amer. Math. Soc.*, 1927.

[6] In terming this process the Heaviside Rule I do not in any sense imply that Heaviside himself would have applied it incorrectly. In fact in one case he adds an extra term which contributes to numerical accuracy although the series itself is

If the operational equation $h = 1/H(p)$ admits of formal series expansion in the form

$$h = a_0 + a_1\sqrt{p} + a_2 p + a_3 p\sqrt{p} + a_4 p^2 + \ldots, \tag{4}$$

a solution, usually divergent and asymptotic, results from discarding the terms in integral powers of p, and replacing $p^n\sqrt{p}$ by $\frac{d^n}{dt^n}\frac{1}{\sqrt{\pi t}}$, whence

$$h \sim a_0 + \left\{ a_1 + a_3\frac{d}{dt} + a_5\frac{d^2}{dt^2} + \cdots \right\}\frac{1}{\sqrt{\pi t}}\cdot \tag{5}$$

As stated in a forthcoming paper, this divergent series is a true asymptotic expansion, as defined by Poincare, if and only if, the singularities in $1/H(p)$ all lie to the left of the imaginary axis in the complex plane. Otherwise the series may require the addition of an extra term or factor, or even be quite meaningless.

An excellent illustration of the preceding principle is furnished by the operational equation,

$$h = \frac{\sqrt{p}}{\sqrt{p+\lambda}}\cdot \tag{6}$$

For convenience and without loss of essential generality we take $|\lambda| = 1$ and $\lambda = e^{i\theta}$; that is, the parameter λ may lie anywhere on a circle of unit radius in the complex plane.

Now the solution of (6) is easily derived by well known processes of the operational calculus: it is

$$h(t) = \frac{1}{\pi}\int_0^t \frac{e^{-\lambda\tau}}{\sqrt{\tau}\sqrt{t-\tau}}d\tau \tag{7}$$

$$= \frac{e^{-\lambda t}}{\pi}\int_0^t \frac{e^{\lambda\tau}}{\sqrt{\tau}\sqrt{t-\tau}}d\tau. \tag{8}$$

The solution is also known to be [7]

$$h(t) = e^{-(\lambda t/2)} I_0\left(\frac{\lambda t}{2}\right), \tag{9}$$

where $I_0(\lambda)$ is the Bessel function $J_0(ix)$.

a true asymptotic expansion. On the other hand Heaviside in his frequent applications of the Rule gives no hint or indication of the restrictions imposed on its applicability. Fortunately in most applications of the operational calculus to physical problems, the Rule leads to correct results.

[7] See formula (p) of the table of integrals in Chap. IV, "Electric Circuit Theory and Operational Calculus."

Now return to the operational equation (6), and expand as follows, without reference to convergence,

$$
\begin{aligned}
h &= \frac{1}{\sqrt{\lambda}}\left(1+\frac{p}{\lambda}\right)^{-1/2}\sqrt{p} \\
&= \frac{1}{\sqrt{\lambda}}\left(1-\frac{1}{2}\left(\frac{p}{\lambda}\right)+\frac{1}{2!}\left(\frac{1}{2}\right)\left(\frac{3}{2}\right)\left(\frac{p}{\lambda}\right)^2-\cdots\right)\sqrt{p}.
\end{aligned}
$$

Application of the Heaviside Rule now gives the divergent solution

$$
\begin{aligned}
h(t) &\sim \left\{1+\frac{1}{1!}\left(\frac{1}{4\lambda t}\right)+\frac{1^2\cdot 3^2}{2!}\left(\frac{1}{4\lambda t}\right)^2+\cdots\right\}\frac{\sqrt{\pi\lambda t}}{1}, \\
&\sim S(\lambda t). \qquad (10)
\end{aligned}
$$

We have now to distinguish three cases:

1. $\lambda_R > 0$. (Real part of $\lambda > 0$.)

In this case it can be shown from (7) that [8]

$$h(t) \sim S(\lambda t) \qquad (11)$$

and that the Heaviside Rule leads to a true asymptotic expansion, as defined by Poincare. When $\lambda = 1$, by the known expansion of the right hand function in equation (9) we find that the error committed by stopping with any term in the divergent series is less than that term. This property, however, does not characterize the series for all complex values of λ for which the real part is positive.

2. $\lambda_R < 0$, $\lambda = -\mu$, $\mu_R > 0$.

In this case, comparison of (8) with (7), gives by aid of (11),

$$h(t) \sim e^{\mu t}S(\mu t), \qquad (12)$$

which again is a true asymptotic expansion. The expansion differs, however, from that given by the Heaviside Rule, by the factor $e^{\mu t}$, and the alternation in sign of the odd terms of the series.

3. $\lambda_R = 0$, $\lambda = i\omega$.

In this case it is easily shown that [9]

$$h(t) = e^{-(i\omega t/2)}J_0\left(\frac{\omega t}{2}\right), \qquad (13)$$

where J_0 is the Bessel function of order zero. From the known asymptotic expansion of this function, we find that

$$h(t) \sim e^{-(i\omega t/2)}\left[e^{(i\omega t/2)}S(i\omega t)\right]_{\text{Real Part}} \qquad (14)$$

with an error less than the last term included.

---

[8] L.c. by the process described in Chap. V.

[9] L.c. formula (n) of table of integrals, Chap. IV.

Perhaps the simplest way of establishing the Heaviside Rule for the asymptotic solution of the operational equation $h = 1/H(p)$ and the conditions under which it is valid, is as follows: We start with the integral equation

$$\int_0^\infty h(t)e^{-pt}dt = 1/pH(p) \tag{15}$$

and specify that the singularities of $1/pH(p)$ and its derivatives are all confined to the left hand side of the complex plane, except at the point $p = 0$, in the neighborhood of which

$$\frac{1}{pH(p)} = \frac{a_0}{\sqrt{p}} + a_1 + a_2\sqrt{p} + a_3 p + a_4 p\sqrt{p} + \cdots. \tag{16}$$

In other words, $1/pH(p)$ admits of expansion in powers of $\sqrt{p}$.

Now since

$$\int_0^\infty \frac{e^{-pt}}{\sqrt{\pi t}}dt = \frac{1}{\sqrt{p}} \tag{17}$$

we have from (15)

$$\int_0^\infty \left(h - \frac{a_0}{\sqrt{\pi t}}\right) e^{-pt}dt = \frac{1}{pH(p)} - \frac{a_0}{\sqrt{p}}. \tag{18}$$

By virtue of the restrictions imposed on $1/pH(p)$, equation (18) is valid at $p = 0$, whence by (16)

$$\int_0^\infty \left(h - \frac{a_0}{\sqrt{\pi t}}\right) dt = a_1. \tag{19}$$

Now differentiate (18) with respect to p; we get

$$\int_0^\infty \left(h - \frac{a_0}{\sqrt{\pi t}}\right) te^{-pt}dt = -\frac{d}{dp}\left(\frac{1}{pH(p)} - \frac{a_0}{\sqrt{p}}\right). \tag{20}$$

Now add $\int_0^\infty \frac{a_2}{2}\frac{e^{-pt}}{\sqrt{\pi t}}dt$ to the left of (20) and its value $a_2/2\sqrt{p}$ to the right hand side; we have

$$\int_0^\infty \left(h - \frac{a_0}{\sqrt{\pi t}} + \frac{a_2}{2t}\frac{1}{\sqrt{\pi t}}\right) te^{-pt}dt = -\frac{d}{dp}\left(\frac{1}{pH(p)} - \frac{a_0}{\sqrt{p}}\right) + \frac{a_2}{2\sqrt{p}}. \tag{21}$$

Now set $p = 0$; from (16) we have

$$\int_0^\infty \left(h - \frac{a_0}{\sqrt{\pi t}} + \frac{a_2}{2t}\frac{1}{\sqrt{\pi t}}\right) t dt = -a_3, \tag{22}$$

a formula which again is valid by reason of the restrictions imposed on $1/pH(p)$.

Proceeding in this manner we get the formula

$$\int_0^\infty (h - S_n)\cdot t^n dt = (-1)^n n! a_{2n+1}, \tag{23}$$

where

$$S_n = \frac{1}{\sqrt{\pi t}}\left(a_0 - \frac{a_2}{2t} + 1.3\frac{a_4}{(2t)^2} + \cdots \right.$$
$$\left. + (-1)^n 1.3 \cdots (2n-1)\frac{a_{2n}}{(2t)^n}\right)$$

$$= \text{first } (n+1) \text{ terms of the divergent Heaviside series.} \tag{24}$$

Also since

$$S_{n+1} = S_n + (-1)^{n+1}\frac{1.3\cdots(2n+1)}{(2t)^{n+1}}\frac{a_{2n+2}}{\sqrt{\pi t}} \tag{25}$$

we have from (23) by changing n to $(n + 1)$,

$$\int_0^\infty \left(h - S_n - (-1)^{n+1}\frac{1.3\cdots(2n+1)}{(2t)^{n+1}}\frac{a_{2n+2}}{\sqrt{\pi t}}\right) t^{n+1} dt$$
$$= (-1)^{n+1}(n+1)! a_{2n+3}. \tag{26}$$

Equations (23) and (26) establish the fact that $(h - S_n)$ converges, for indefinitely great values of t, at least as rapidly as $1/t^{n+1}\sqrt{t}$, since otherwise the integrand of (26) would diverge; stated in mathematical notation

$$h - S_n = 0(1/t^{n+3/2}). \tag{27}$$

Consequently the series S when divergent is a true asymptotic expansion, as defined by Poincare, of the function h.

The foregoing says nothing, it will be noted, regarding the error committed when S_n is employed to compute the function h. Nothing, in general, can be said about this question, which requires an independent investigation in every specific problem. In some cases the error will be less than the magnitude of the last term of S_n, but this is the exception rather than the rule. In other exceptional cases the series may even be absolutely convergent.

The foregoing results can undoubtedly be derived by integration of the Bromwich integral (2) along the contour suggested by March (*l.c.*). Wiener in his paper on "The Operational Calculus" (*Math. Annalen*, Bd. 95, 1925) gives an entirely different treatment of the problem. The operational calculus he deals with, however, differs under some circumstances from that of Heaviside, as Wiener himself remarks. A paper by Tibor v. Stacho on "Operatoren Kakül von Heaviside und Laplaceshe Transformation" (publication 1927 VI 15 by the Hungarian University, Francis Joseph) may also be consulted.

II

Subject to certain well known restrictions a function $f(t)$ can be expressed as the Fourier integral

$$f(t) = \frac{1}{2\pi i}\int_{-i\infty}^{i\infty} F(p)e^{pt}dp. \tag{28}$$

the path of integration being along the imaginary axis. We assume for the moment that this equation is valid.

Now suppose that $f(t)$ represents a force applied to an electrical or dynamic system whose "steady state" or forced response to an applied force $F(p)e^{pt}$ is

$$\frac{F(p)}{H(p)}e^{pt}.$$

Then the *forced* response $g(t)$ of the system to the applied force $f(t)$ is given by

$$g(t) = \frac{1}{2\pi i}\int_{-i\infty}^{i\infty} \frac{F(p)}{H(p)}e^{tp}dp. \tag{29}$$

However, in applying the foregoing to the Heaviside problem we encounter an initial difficulty. This is that if $f(t)$ is taken as the unit function (zero before unity after, $t = 0$) it does not admit of formulation as the Fourier integral (28). The unit function, however, when multiplied by e^{-ct} when c is a positive real constant, does admit of such formulation, and it is easy to show that the unit function itself is given by

$$\frac{1}{2\pi i}\int_{c-i\infty}^{c+i\infty} \frac{e^{pt}}{p}dp \qquad c > 0. \tag{30}$$

Consequently, if the unit function is the force impressed on the system, the *forced* response is

$$k(t) = \frac{1}{2\pi i}\int_{c-i\infty}^{c+i\infty} \frac{e^{pt}}{pH(p)}dp \qquad c > 0. \tag{31}$$

If now all the singularities of the integrand lie to the left of the imaginary axis, then $k(t) = h(t)$ and (31) is the formulation of the Heaviside problem. Suppose, however, that the electrical or dynamic system specified by $H(p)$ is "unstable"; that is, it contains some internal source of energy which makes its transient oscillations increase with time t instead of dying away. In such a case $H(p)$ will have zeros to the right of the imaginary axis, and in order that (31) shall be the solution of the Heaviside problem, c must be taken so large that all the singularities of the integrand lie to the left of the path of integration. Consequently

$$h(t) = \frac{1}{2\pi i}\int_{c-i\infty}^{c+i\infty} \frac{e^{pt}}{pH(p)}\,dp, \tag{2}$$

provided c is so chosen that all the singularities lie to the left of the path of integration in the complex plane. This is Bromwich's formulation of the Heaviside problem.[10]

From the foregoing it follows that the Fourier integral

$$\frac{1}{2\pi i}\int_{-i\infty}^{i\infty} \frac{e^{pt}}{pH(p)}\,dp \tag{2a}$$

is, in general, the formulation of the Heaviside problem if and only if, all the singularities of the integrand lie to the left of the imaginary axis. If there are singularities on the imaginary axis, the integral is ambiguous, while if there are singularities to the right of the imaginary axis, the integral gives an incorrect solution of the Heaviside problem.[11]

As a simple example consider the operational equation

$$h = 1/H(p) = \frac{p}{p-\beta},$$

where the real part β_r of β is positive. The correct solution as given by either (2) or (3) is

$$\begin{aligned} h &= 0 \qquad && t < 0 \\ &= e^{\beta t} && t > 0, \end{aligned}$$

---

[10] The appropriate mathematical methods of solving the infinite integral (2) are dealt with in great detail by Jeffreys in his "Operational Methods in Mathematical Physics" (Cambridge University Tracts).

[11] To prevent misunderstanding it should be stated that the application, when permissible, of the classical Fourier integral (2a) to the Heaviside problem, was known long prior to the work of Bromwich. Bromwich's essential and important contribution lay in showing that the path of integration must be shifted to the right of all the singularities, together with a verification of an important form of solution, first given by Heaviside, of the operational equation.

whereas the Fourier integral (2a) gives

$$\begin{aligned} h &= -e^{\beta t} & \quad t < 0. \\ &= 0 & \quad t > 0. \end{aligned}$$

There is another reason why care must be exercised in applying the classical Fourier integral to the Heaviside problem. This is that in solving the operational equation, $h = 1/H(p)$, the appropriate expansion of $1/H(p)$ may introduce singularities on or to the right of the imaginary axis in the component terms. This offers no difficulty if either (2) or (3) is employed, but renders the Fourier integral (2a) inapplicable. As an example consider the equation

$$h = \frac{1}{\sqrt{p} + 1}.$$

One form of solution is gotten by multiplying numerator and denominator by $\sqrt{p} - 1$, whence

$$h = \frac{\sqrt{p}}{p - 1} - \frac{1}{p - 1}$$

and each term has a singularity at $p = 1$.

A physical interpretation of the foregoing may not be without interest. Suppose that an elementary force $F(p)e^{pt}dp$, where $p = c + i\omega$, is applied at an indefinitely remote past (negative) time to a system specified by $H(p)$. The response of the system is then

$$\frac{F(p)}{H(p)} e^{pt}dp + T_p(t)dp,$$

where $T_p(t)dp$ is the concomitant transient or characteristic oscillation of the system. If c is chosen sufficiently large then at least for $t \geq 0$ the transient term can be made as small as we please compared with the first term. Finally if the impressed force is the unit function (zero before, unity after, time $t = 0$) and it is written as

$$\frac{1}{2\pi i} \int_{c-i\infty}^{c+i\infty} \frac{e^{tp}}{p} dp,$$

the total response and therefore $h(t)$ is given by

$$\frac{1}{2\pi i} \int_{c-i\infty}^{c+i\infty} \frac{e^{tp}}{pH(p)} dp,$$

provided c is sufficiently large to make the transient term √√

$$\int_{c-i\infty}^{c+i\infty} T_p(t)dp$$

negligibly small. Analytically this requires that c be so large that the zeros of $pH(p)$ shall all lie to the left of the axis $p_R = c$.

III

The foregoing discussion tacitly assumes the existence of an unique solution of the operational equation. On the part of the physicist this assumption is entirely proper because if the operational equation is the symbolic formulation of a correctly set physical problem an unique solution must and does exist. When approached from the purely mathematical standpoint, however, the case is different and there is no assurance of the existence of a solution. As an example consider the operational equation

$$h = e^p$$

The corresponding integral equation

$$\frac{e^p}{p} = \int_0^\infty h(t)e^{-pt}dt \qquad p_R > 0$$

has no solution, while Bromwich's formula

$$h(t) = \frac{1}{2\pi i}\int_{c-i\infty}^{c+i\infty} \frac{e^p}{p} e^{tp} dp$$

gives

$$\begin{aligned} h &= 0 \qquad & t &< -1 \\ &= 1 \qquad & t &> -1 \end{aligned}$$

which is obviously incorrect. As a matter of fact the operational equation itself has no solution.

To formulate the necessary and sufficient conditions for the existence of a solution we may proceed as follows: If a solution exists it is given by either of the equations

$$h(t) = \frac{1}{2\pi i}\int_{c-i\infty}^{c+i\infty} f(p)e^{tp}dp, \tag{2}$$

$$f(p) = \int_0^\infty h(t)e^{-pt}dt \qquad p_R \geq c, \tag{3}$$

where $f(p)$ denotes $1/pH(p)$. Substitution of the value of $h(t)$, as given by (2), in (3), gives the transform

$$f(p) = \frac{1}{2\pi i}\int_0^{\infty} e^{-pt}dt \int_{c-i\infty}^{c+i\infty} f(z)e^{tz}dz. \tag{32}$$

In addition, since $h(t) = 0$ for $t < 0$, we must have

$$\frac{1}{2\pi i}\int_{c-i\infty}^{c+i\infty} f(p)e^{tp}dp = 0 \text{ when } t < 0. \tag{33}$$

Equations (32) and (33) formulate the necessary and sufficient restrictions on $f(p)$ for the existence of a solution of the operational equation

$$h = pf(p) = 1/H(p).$$

To correlate the transform (32) more closely with the classical Fourier transform, write $p = u + i\omega$ and

$$f(u + i\omega) = \phi(\omega) \qquad u \text{ and } \omega \text{ real}.$$

Then the transform (32) becomes

$$\phi(\omega) = \frac{1}{2\pi}\int_0^{\infty} e^{-i\omega t}dt \int_{-\infty}^{\infty} \phi(x)e^{itx}dx \tag{34}$$

for all values of $u \geq c$. Also since $h(t) = 0$, for $t < 0$, the lower limit of integration with respect to t in (33) may be replaced by $-\infty$, whence

$$\phi(\omega) = \frac{1}{2\pi}\int_{-\infty}^{\infty} e^{-i\omega t}dt \int_{-\infty}^{\infty} \phi(x)e^{itx}dx, \tag{35}$$

which is the classical Fourier transform.

The foregoing naturally suggests a few remarks regarding the mode of approach to the operational calculus. If we regard, as Heaviside certainly did, the operational equation as the symbolic formulation of a definite physical problem, it is not permissible to define the significance of the operator p *a priori*. The meaning of the operator p and methods of solution of the equation must be so determined as to give the correct solution of the original physical problem. Heaviside's procedure here was purely heuristic and "experimental"; equations (2) and (3), however, provide a sound logical basis for the development of the operational calculus. On the other hand, from the purely mathematical standpoint it is possible to develop an operational calculus on the basis of certain mutually consistent definitions and conventions adopted at the outset, just as it is possible to develop different geometries and algebras. An operational calculus so developed, however, may or may not agree with that of Heaviside and may or may not give the correct solution of the Heaviside problem. In a number of recent papers on the Heaviside operator this procedure has been adopted. To the writer this appears both illogical and doubtful, and is certainly not the method of Heaviside himself, as is sometimes implied.

In the interpretation of the operational equation $h = 1/H(p)$ it is, in the writer's opinion, extremely important to recognize the fact that it is not a true equation and has no literal significance of itself, but is simply and solely the symbolic or shorthand way of writing down equation (2) or its equivalent (3). If this fact is kept clearly in mind the 'operator' p loses the mysterious character it seems to possess for so many students and all real danger of misinterpretation and incorrect solution is eliminated. In the writer's opinion, Heaviside's achievement in the development of his operational calculus does not consist in inventing a novel and mysterious kind of mathematics, but in formulating a body of rules and processes whereby recourse to the actual equations of the problem is rendered unnecessary.

There is another fact which it is also important to clearly recognize. In the original differential equations from which the operational equation is derived, the symbol p^n denotes d^n/dt^n and its reciprocal p^{-n}, corresponding multiple integration, and the index n is always integral. If, as in the case in important electrotechnical problems, non-integral or fractional powers of the symbol p occur in the operational equation, it is due to algebraic manipulations and operations, which in essence rob p of its original significance. That is to say, in such cases it is not permissible nor indeed possible to assign to the operator p its original significance. For example the operational equation

$$h = \sqrt{p}$$

does not mean

$$h(t) = \left(\frac{d}{dt}\right)^{1/2} \cdot 1 \qquad (1 = \text{unit function})$$

which is itself meaningless, but simply

$$h(t) = \frac{1}{2\pi i}\int_{c-i\infty}^{c+i\infty} \frac{e^{tp}}{\sqrt{p}}\,dt \qquad c > 0$$

or

$$\frac{1}{\sqrt{p}} = \int_0^\infty h(t)e^{-pt}dt \qquad p_R > 0$$

More broadly stated, the operational equation is the shorthand statement of true equations in which p has lost its original significance and is simply the complex argument of functions which obey all the laws of algebra and analysis.

Failure to recognize these simple principles is responsible for a large amount of confusion, loose reasoning and profitless discussion of so called 'fractional differentiation,' a term which, to the writer at least, is quite meaningless. On the other hand, their recognition should go far towards removing whatever mystery may have surrounded the Heaviside operator and the Heaviside processes.

CORRECTION SLIP FOR ISSUE OF JANUARY, 1930

Page 153: Equation (10) should read

$$h(t) \sim \left\{ 1 + \frac{1}{1!}\left(\frac{1}{4\lambda t}\right) + \frac{1^2 \cdot 3^2}{2!}\left(\frac{1}{4\lambda t}\right)^2 + \cdots \right\} \frac{1}{\sqrt{\pi \lambda t}},$$

$$h(t) \sim S(\lambda t) \tag{10}$$

Contemporary Advances in Physics, XIX.

Fusion of Wave and Corpuscle Theories.

By KARL K. DARROW.

In this article certain of the simple and familiar phenomena of optics and of electronics—for instance, refraction at a boundary between two media, and diffraction by a grating—are interpreted by *both* of the theories, undulatory and corpuscular, which have so often been condemned as incompatible with one another; the attitude being, that the theories may be brought into concordance by modifying one at least in ways which, extraordinary as they seem, do not quite destroy its character.

NOT quite five years ago I published in this journal an article entitled *Waves and Quanta*, expounding there the data which invited a corpuscular theory of light, regardless of the great array of classical phenomena of optics which demanded with no less insistence the long-triumphant undulatory theory. Today, not only are those data still extant and undeniable; they have been reinforced by observations on electron-streams which have compelled a wave-theory of free negative electricity, despite the very abundant evidence for free corpuscular electrons. Most physicists expect that not only light and negative electricity, but whatever other fundamentals there may be—meaning, probably, positive electricity and nothing else—will be found to conform in some ways to simple wave-theory, and in some to simple particle-theory. Most physicists, I think, would concede that the two ideas must be forced into one scheme, whatever violence it may entail to others of our preconceptions, inborn or inbred. We must stretch the theories and our minds, so that corpuscles and waves shall appear no longer as alternatives of which election must be made, but as complementary aspects of one reality.

To make a beginning with this process of stretching, I propose to treat some of the very simplest and most familiar of the phenomena, which up to lately have been interpreted by *one only* of the theories: phenomena such as the refraction of light in passing from air to water, the bending of the paths of electrons in passing from vacuum into metal, the diffraction of light and electrons from a ruled diffraction-grating. (None of these examples, incidentally, involves a theory of the structure of the atom.) Each of them shall be interpreted by the *other* theory—not in order to substitute the *other* for the *one*, but in order to practice the art of using *both* theories in alliance.

REFRACTION OF WAVES AND REFRACTION OF CORPUSCLES.

I presume that every textbook of optics and every history of physics informs its readers that anciently there was a controversy between a wave-theory of light (attributed to Huyghens) and a corpuscular theory (accredited to Newton) which was totally decided in 1850 by an experiment of Foucault. Light is refracted toward the normal in passing from air to water, and should therefore move more rapidly in water than in air if it consists of particles, but not so rapidly if it consists of waves—so runs the argument. Foucault and Fizeau discovered that light does move less rapidly in water than in air.[1] Let us analyze the argument more closely before deciding what was proved.

The reasoning from the "wave-theory" is usually made in graphic fashion by showing "Huyghens' construction" (Fig. 1) which should remind many a reader of his high school days! This is a very crude form of wave-theory, much too primitive to account for most of the phenomena which the physicist has in mind when he says that light (or electricity, or matter) is of the nature of waves; but for the present purpose it will do.

In Fig. 1, AA' is the trace, on the plane of the paper, of a wavefront moving through air (say) in the direction LM toward the boundary between air and water. It is the trace of the wavefront at a particular moment, say t; at a later moment, say t', the front has moved on to another position BB'. Denote by v the speed of the wavefront in air; then the perpendicular distance between BB' and AA' is equal to $v(t' - t)$. While the wave is advancing through this distance, its intersection with the boundary of the water sweeps over the distance AB, which we will denote by D. Designate by θ the angle between wavefront and boundary, the "angle of incidence." From the diagram one sees immediately:

$$\sin \theta = v(t' - t)/D. \tag{1}$$

Now in Huyghens' view, whenever the oncoming wavefront passed over an atom in the boundary-surface it incited that atom to emit a "wavelet." The circles drawn around various points on the line AB are the traces on the plane of the paper, of halves of those spherical wavelets—the halves expanding downwards into the water. Accord-

---

[1] Foucault usually gets all the credit, but Fizeau and Bréguet were working at the same time, incited by the same suggestion of Arago, and using the same method with differences in detail; and they announced their result only six weeks later. Indeed, at the meeting of the Académie des Sciences (May 6, 1850) at which Foucault reported his success, Fizeau said that if the sun had shone that day or the day before, they too would have had data to present.

ing to "Huyghens' Principle" the ongoing wavefront in the water is the envelope of these spheres. In Fig. 1 they and the ongoing wavefront are represented for the moment t' when the wave in the air reaches B. The radius AC of the wavelet expanding from A is then the distance which light traverses in water during time $(t' - t)$, for that wavelet started when the wave in the air reached A. Denote by v' the speed of light in water and by θ' the angle between the new wavefront and the boundary, the "angle of refraction"; then from the diagram:

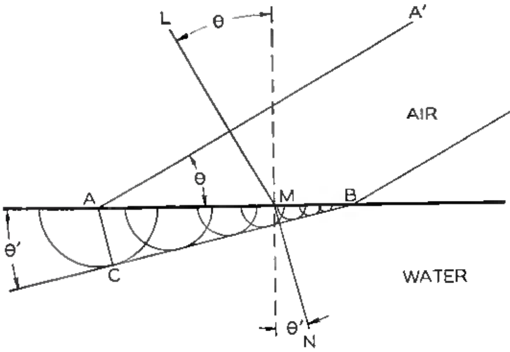


Fig. 1.

$$\sin \theta' = v'(t' - t)/D \tag{2}$$

and from (1) and (2) together, we obtain:

$$\sin \theta/\sin \theta' = v/v'. \tag{3}$$

From this familiar equation it follows in general, that the ratio $(\sin \theta/\sin \theta')$ is independent of the angle of incidence. (It is called the *index of refraction* of the second medium with respect to the first; I denote it hereafter by N.) Also it follows in particular, that when light is refracted towards the normal the wavefronts must move more slowly in the second medium than in the first, which is what Foucault verified, or rather, thought he had verified.

Now try it by the corpuscle-theory. In Fig. 1, I have the line LMN redrawn as a heavy line, and the lines at right angles to it left out; for the line LMN, one of the "rays" of light, is now to be interpreted as the path of a corpuscle, and there are no wavefronts.

So long as the corpuscle is too far from the boundary-surface to feel any force from the water, it moves in a straight line with unchanging momentum; for the forces exerted on it by the air, being equally applied in all directions, balance one another out. In the region near the boundary, this remains the truth for the components of force parallel to the surface; but the components along the normal, applied respectively from the direction towards the air and the direction towards the water, need not be perfectly equal. After the corpuscle has gone through the transition region and reached the depths of the water, it continues in a straight line with a momentum of which the component parallel to the boundary—the "tangential" component, say—is still the same as it was in the air, while the normal component is changed. Denote by p_t and p_n these two components of the original momentum of the particle through the air, by p the magnitude of their resultant which is the original momentum; by p_t', p_n' and p' the corresponding quantities for the final flight of the corpuscle through the water. From Fig. 2 we see:

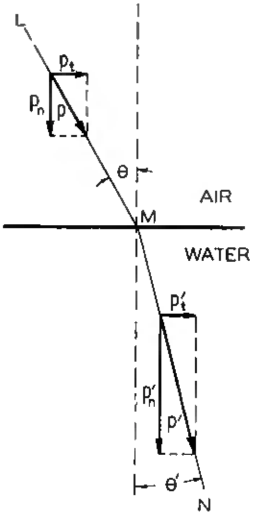


Fig. 2.

$$\sin\theta = p_t/\sqrt{p_t^2 + p_n^2} = p_t/p, \ \sin\theta' = p_t'/p, \tag{4}$$

and since $p_t = p_t'$:

$$\sin\theta/\sin\theta' = p'/p. \tag{5}$$

The corpuscle-theory therefore leads to the statement that the sines of the angles of incidence and refraction stand to one another as the momenta of the corpuscle in the first medium and in the second; and when light is refracted towards the normal, the corpuscles must move with a greater momentum in the second medium than in the first.

Comparing the equations (5) and (3) to which the two conceptions lead, one sees that far from contradicting one another, they are both acceptable, provided that:

$$p/p' = v'/v. \tag{6}$$

We may hold both the theories simultaneously, we may interchange the two at will, provided we assume that the momentum of the corpuscles varies inversely as the speed of the wavefronts. In spite of the outcome of Foucault's experiment, we may adopt either the wave-theory or the corpuscle-theory or both at once to describe refraction, provided we assume that when a beam of light is refracted toward the normal, the speed of the wavefronts diminishes but the momentum of the corpuscles grows greater.

Why then did everyone concede that the corpuscular theory of light was killed by the experiment of Foucault? Because everyone was making two assumptions which seemed so obvious as to be hardly worth the stating, and so certain that it would have been regarded as absurd to call either into question:

(*A*) It was being assumed, that the momentum of a corpuscle must always be strictly proportional to its velocity; in other words, that the mass of a corpuscle must be invariant.

(*B*) It was being taken for granted that in a wave-theory of light the speed of the waves, and in a corpuscle-theory of light the speed of the corpuscles, must be identified with the actual speed of light as measured in any actual experiment.

When these assumptions are made, equation (5) goes over into the form,

$$\sin\theta/\sin\theta' = p'/p = v'/v, \tag{7}$$

which is contradictory to equation (3) and disproved by the experiment of Foucault.

But it no longer seems radical to change the first of these assumptions, for it is known from observation that there are particles—electrons, for example—of which the mass is not invariant, but depends upon the speed. For such a particle the momentum is not exactly proportional to the velocity. It is then not quite so revolutionary to go further, and suppose that the corpuscle of light is of so strange a nature that its velocity and its momentum are in magnitude inversely proportional to one another. If one made this supposition then one could accept the second assumption, and still explain the refraction of light by the corpuscle-theory.

Even the second assumption, however, is not sacred. It may seem absurd to set up a wave-theory of light, and then say that the speed of the wavefronts is not to be identified with the measured speed of light. It does seem absurd to set up a corpuscle-theory, and then say that the speed of the corpuscles is not necessarily the same as that of light. Yet it may turn out in the end that a theory of either kind is strengthened, and made more competent to account for a variety of facts, by abandoning that easy and natural identification. I will try to prove by actual examples that it does so turn out. Meanwhile I summarize this section in a sentence:

If we wish to interpret light, or electricity, or matter, by both a corpuscle-theory and a wave-theory, the momentum of the corpuscles must be supposed to vary inversely as the speed of the waves.

I have omitted the special reference to refraction, for any more general theory must include that particular case, or fall down completely; I have added allusions to electricity and matter, for the test of any alteration of the two classical assumptions will depend chiefly on whether it helps in understanding the wavelike properties of these two, and not of light alone.

We now carry the wave-theory a great step beyond the primitive form in which Huyghens left it, by introducing the ideas of *frequency and wave-length.*

Wave-length of Waves and Momentum of Corpuscles

Instead of the single "wavefront" of Fig. 1, suppose a train of sine-waves of frequency ν, period $T(= 1/\nu)$, wave-length λ and wave number $\mu(= 1/\lambda)$ travelling through air along the course LMN. For definiteness, think of sound-waves. The condensation[1] of the air conforms to the equation:

$$\rho = \rho_0 \sin 2\pi\,(\nu t - \mu s + \alpha), \tag{8}$$

wherein s stands for distance measured from some arbitrary plane perpendicular to LM, and α for some constant. I write the equation down because one like it (or more than one) occurs in every wave-theory. In that of light there are six such equations, with components of electric and magnetic field strength replacing ρ; but it will be sufficient to think of one. In the wave-theory of matter there is one, with a quantity of very abstract meaning replacing ρ.

Now when the wave train passes through into the water, its frequency remains the same. With sound-waves, or any mechanical vibrations of matter, this is obvious; two pieces of matter in continuous contact must vibrate in unison, or not at all. We generalize this statement to cover light-waves, and waves of other varieties later

---

[1] The excess of the density over the normal value, divided by the normal value.

to be considered. Using primes to designate the values which things have in the second medium, we put:

$$\nu' = \nu. \tag{9}$$

The speed of the waves is the product of their wave-length by their frequency:

$$v = \nu\lambda, \qquad v' = \nu'\lambda'; \tag{10}$$

consequently:

$$v'/v = \lambda'/\lambda. \tag{11}$$

The wave-lengths of the wave train on the two sides of the boundary vary directly as the speeds.

Return now to the last section, and introduce this result into equation (6); one gets:

$$p'/p = \lambda/\lambda' \tag{12}$$

which means: we can interpret refraction of light (or of electricity, or of matter) by both the wave-theory and the corpuscle-theory, provided that we make the momentum of the corpuscle vary inversely as the wave-length of the waves.

Write accordingly,

$$p\lambda = \text{constant}. \tag{13}$$

Now there are several remarkable experiments which show that this relation actually holds, and moreover that the constant which appears in it is the universal constant h of Planck:

$$p = h/\lambda. \tag{14}$$

For instance, one may pour a stream of X-rays—that is to say, high-frequency light—into a gas, after having measured its wave-length in the known and reliable way depending on one of the phenomena in which X-rays behave as waves. A certain portion of the rays is scattered; it is scattered as though it consisted of corpuscles, each of which strikes an individual free electron and bounces off, the electron meanwhile recoiling from the blow.[2] Further analysis of the data shows that there is conservation of momentum—that the momentum which the electron gains is equal to that which the corpuscle of light has lost, *provided that the momentum of this latter is equal to the quotient of h by the wave-length of the rays.* For the wave-length of the scattered X-rays, measured in the same way as that of the primary rays was measured, is not the same as theirs; and the difference between the values of h/λ, before and after scattering, is equal to the momentum which the electron received.

[2] The Compton effect (cf. the seventh article of this series).

Again, one may pour a stream of electrons against a crystal or an optical ruled grating, after having measured the speed of the electrons in one of the well-known ways depending ultimately on the deflection of such a beam in known electric and magnetic fields.[3] The mass of the electrons being known, one knows also their momentum. Now the crystal or the grating, whichever it may be, forms from the primary beam a diffraction-pattern of new beams. Well! the formation of a diffraction-pattern is the primary reason for saying that light is wavelike, and it gives the primary way of measuring wave-length of light. One is equally obliged to admit that a stream of free negative electricity is wavelike, and to accept the value for its wave-length which the diffraction-pattern gives. Again it turns out that the wave-length is equal to the quotient of h by the momentum of the electrons.

It may be objected that in all of those experiments, the corpuscles were observed in a vacuum. Compton measured X-rays before and after scattering, but during the measurements they were in vacuum or at any rate in air. Davisson and Germer, Thomson and Rupp, observed electrons returning through the same evacuated space as they had crossed on their way to the diffracting lattice. One might emphasize that all these savants compared momenta and wave-lengths for different beams in the same medium instead of comparing them for the same beam in different media. The distinction is certainly worth noticing; but happily there are experiments which bear directly on refraction. Davisson and Germer measured, not precisely the refraction of an electron-stream passing from vacuum into nickel, but a minor perturbation of the diffraction-pattern which is due to that refraction. We will analyze their result, for nothing shows more clearly the relations—or lack of relation, the reader may think—between speed of waves, speed of corpuscles and measured speed of stream.

Davisson and Germer came to values of the index of refraction ($\sin\theta/\sin\theta'$) which were greater than unity—which corresponded therefore to a bending of the stream towards the normal, as it passed from vacuum into nickel—which therefore signified that the speed of the waves is not so great in nickel as in air.

On the other hand, it is known that when an individual electron passes from vacuum into a metal, its kinetic energy and its velocity increase as it goes through the surface. We have in fact the situation described in the corpuscle-theory picture of refraction, a few pages back. Return to equations (4) and (5), and consider a corpuscle for

---

[3] The experiments of Davisson and Germer, of G. P. Thomson, and of Rupp (cf. the eighteenth article of this series).

which the momentum p, the velocity u, the kinetic energy K, the mass m are related to one another as in Newtonian mechanics—properties which are practically those of electrons except when these are moving much more rapidly than any involved in these experiments:

$$p = mu, \qquad K = \tfrac{1}{2}mu^2. \tag{15}$$

Use u_t and u_n to denote tangential and normal components of speed; use primes to designate the values which things have in the second medium (nickel). Starting from equation (5), we continue:

$$\sin\theta/\sin\theta' = N = M'/M = u'/u;$$
$$N^2 - 1 = (u'^2 - u^2)/u^2 = (K' - K)/K. \tag{16}$$

The quantity $(K' - K)$ is the gain in kinetic energy which the electron wins on passing into the nickel; and this gain, as I have said, is positive; hence by equation (16) the index of refraction must be greater than unity. This is in agreement with the result of Davisson and Germer; the agreement, in fact, appears to be quantitative.[4]

It is always pleasant to get an agreement; but note how we got this one. We got it by dropping the assumption that the speed of the corpuscles and the speed of the waves must be the same. Or rather, by not making that assumption. For though the fact of experience is always the same—the swerving of the electron-stream *toward* the normal as it enters the nickel—it is interpreted by the two theories in opposite ways; the waves are slowed down, but the corpuscles are speeded up, in passing from the vacuum to the metal. Even if wave-speed and corpuscle-speed were the same in empty space, they could not be the same in any other medium.

This is more serious than it may appear at first. It amounts in effect to saying that a beam of free negative electricity has two different speeds; one when we visualize it as a jet of particles, another quite different when we visualize it as a train of waves.

But is not one of these "the right one" and the other "a wrong one," and can we not settle between them by measuring the actual time which the electricity takes to pass a measured distance? Let us examine this possibility. We shall find that after all it is not so easy to evade the ambiguity in such a fashion.

PHASE-SPEED AND GROUP-SPEED

Suppose an endless train of perfect monochromatic sine-waves marching along through space. For definiteness, think again of sound-

---

[4] There is a remarkably interesting correlation between these results and the new statistical theory of the electron-gas inside the metal (cf. my article in the October 1929 number of this *Journal*, pp. 710–716).

waves. It might seem as if we could measure their speed by picking out one crest, as A of Fig. 3, and checking off with a stop-watch the moments when it passes two fixed markers placed a known distance apart. Not so; for we cannot see or hear or in any way perceive the individual crests. The wave train produces a perfectly uniform tone in the ear which it strikes. If two listeners are stationed at different points along the path of the sound, neither can recognize the moment at which any particular crest glided by. All they can recognize, all they can compare, is the moment of passage of a *perturbation* of the wave train; a sudden beginning, a sudden ending, a transient swelling of the sound. Most measurements of the speed of sound, in fact, are measures of the speeds of something violent—the crack of a pistol or an electric spark, the roar of an explosion—something very unlike a uniform train of sine-waves.[5]

Now a sine-wave with a perturbation is in effect a sum of two or more sine-waves each of endless extent and constant amplitude, but having different wave-lengths and different amplitudes. This statement is the content of Fourier's principle from which the method of Fourier analysis is derived. One might represent even the sudden and violent pulsation of air due to an explosion, or the electrical spasm due to an outburst of static, by a summation of properly-chosen endless monochromatic sine-wave trains. I take however the simplest conceivable case: the wave train composed of only two sine-waves of different wave-lengths.

The reader will probably recall that when the difference between the wave-lengths is only a small fraction of either, this composite wave train resembles a sine-wave with regular fluctuations of amplitude—that is to say, with "beats" (Fig. 3). The maximum or centre of a beat occurs where a crest of one sine-wave coincides with a crest of the other—the minimum between beats, where crest falls together with trough. Denote the two wave-lengths by λ and $\lambda + \Delta\lambda$. One sees by inspection that a wave-length is the same fraction of the distance D between two consecutive beat-maxima, as the discrepancy $\Delta\lambda$ is of the wave-length:[6]

$$D/\lambda = \lambda/\Delta\lambda. \tag{17}$$

Of course this statement is exactly true only in the limit of vanishingly small $\Delta\lambda$. We shall always stay close to this limit, though some of the following statements would be valid even otherwise.

---

[5] I except so-called measurements of the velocity of sound which are really measures of frequency and wave-length in stationary wave-patterns, these being then multiplied together.

[6] The principle of the vernier.

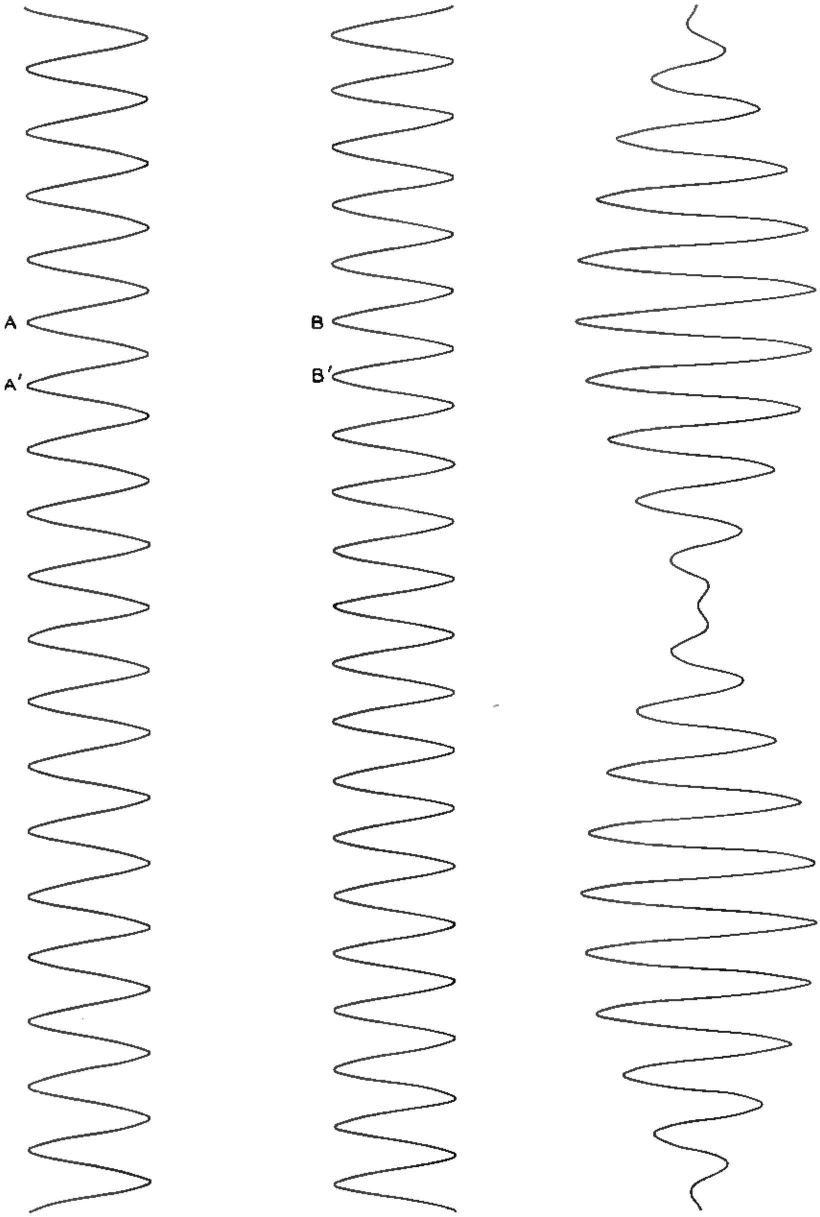


Fig. 3.

Now if the two component waves advance with equal speed, the beats are simply carried along with a speed equal to theirs. But if the velocities of the two component waves are not the same, then the velocity of the beats is not the same as either, nor the mean thereof. It is in fact something totally different.

To see this, imagine that you are moving along with one of the sine-waves; for definiteness, that you are riding on the crest B of the train with the shorter waves (Fig. 3). At a certain moment, say $t = 0$, it coincides with a crest A of the other sine-wave, and you are at the top of the beat. Meanwhile the other train is moving relatively to the first; for definiteness suppose that the longer waves move faster, so that relatively to the shorter they are gliding upward. After a certain time they have gained on the shorter waves by a distance $\Delta\lambda$, the difference between the two wave-lengths. But when this time has elapsed, the top of the beat is no longer where you are, but where the crest B' of the first train coincides with the crest A' of the second. It has dropped back through the distance λ, while the second wave train was getting ahead by the distance $\Delta\lambda$. Perhaps it will be easier to realize that while the second wave train is gaining on the first by λ, the beat is dropping back by the distance D between consecutive beats; by equation (17) this comes to the same thing.

Therefore when the longer waves travel faster than the shorter, the beats travel more slowly than either. If the longer waves were the slower, the beats would travel more rapidly; but this case is never realized in nature, not at least with light-waves[7] and waves of electricity and matter.

We now deduce the formula for the actual value of the speed of the beats. Denote by v and $v + \Delta v$ the speeds of the two sine-waves of which the wave-lengths are λ and $\Delta + \Delta\lambda$, respectively; by g the speed of the beats. It is sufficient to put into notation what has just been said in words. Relatively to the former wave train, the velocity of the latter wave train is Δv, that of the beats is $(g - v)$. Relatively to the former wave train, the latter moves a distance $\Delta\lambda$ while the beats are moving a distance λ in the opposite sense, therefore with a minus sign. Hence:

$$(g - v)/\Delta v = -\lambda/\Delta\lambda \qquad (18)$$

---

[7] The exception to this statement—the case of light having wave-lengths lying within a region of anomalous dispersion of the transmitting substance—has been analyzed by Sommerfeld and L. Brillouin (*Ann. d. Phys.* **44,** pp. 177-202, 203–240; 1914) who find that in this case the group-speed defined by (20) loses its physical importance, and a segment of a wave train is transmitted with a speed never exceeding the speed of light in vacuum. This appears to be related to the absorption which always goes with anomalous dispersion.

and solving for g,

$$g = v - \lambda \frac{\Delta v}{\Delta \lambda}, \tag{19}$$

or going over to the differential notation, which will not only look more natural but will signify that the result which we have just attained is strictly valid in the limit for infinitesimal differences of wave-length:

$$g = v - \lambda(dv/d\lambda). \tag{20}$$

This is the formula for the *group-speed;* for the term "group-speed" is the usual one for what I have been calling "speed of beats." Likewise *phase-speed* is commonly used to denote the speed of the individual sine-wave trains.

The term "group-speed" is in one respect unfortunate; for it implies that any "group," that is to say any sequence of uneven and irregular wave-crests and troughs, is propagated with a perfectly definite speed. However this is true only for the simplified group which we have been considering, the beat formed of no more than two wave trains; and even for this it is exactly true only in the limit, where the wave-length-difference between the trains approaches zero. All other groups change in form as they advance. Now there is always something arbitrary in defining "speed" for something which changes as it goes, like a puff of smoke or a cloud. The arbitrariness is nil in only the limiting case which I have just been formulating. However, it must not be exaggerated. A bunch of irregular crests and troughs may retain enough of its form and compactness, as it travels over a distance many times as great as its width, to justify the statement that it has a speed of its own. And if such a group turns out, on being analyzed in Fourier's way, to consist mainly of sine-waves clustered in a small range of wave-lengths, then its speed will not be far from the value of g computed by equation (20) for a wave-length in that range.

Now these deductions explain a very remarkable experiment by Michelson, which otherwise might have disproved—indeed I do not see how it could have been interpreted otherwise than as destroying—*both* the wave and the corpuscle theory of light. I will preface the account of this experiment by saying that for light in empty space the speed of all wave-lengths is the same,[8] so that there never is any dif-

---

[8] The chief evidence for this statement is astronomical. If light of one color traveled faster than light of another, a luminous star emerging from behind a dark one would be seen first in the faster-travelling hue; in fact there would be a sequence of colors, the same for every emergence of every such star, and spread out over a time-interval proportional to the distance of the stars. Nothing of the sort has ever been observed, although there are plenty of luminous stars revolving around dark ones which regularly occult them.

ference between velocity of groups and velocity of wave-crests; they both have the same universal constant value c. However this cannot be true for light in transparent material media such as glass, water, or carbon bisulphide; for the refractive index of all these media varies from one wave-length to another—they are said to be *dispersive*.

Now Michelson measured the time taken by a flash of light to cover a measured distance, first through air (very nearly the same as vacuum) then partly through air and partly through carbon bisulphide. The source of light shines continuously, and an incessant beam falls on a revolving mirror and is reflected in a continuously-changing direction; a second, stationary mirror receives this reflected beam during a very small fraction of each complete revolution and sends it back, so that the twice-reflected beam is a series of segments cut from the primary beam. It was the time taken by the segments to travel a known distance which Michelson measured.[9] Reasoning back from the data, he computed that they took (1.76 ± 0.02) times as long to go a given distance in carbon bisulphide as in air. But the refractive index of carbon bisulphide, in the range of the spectrum where Michelson's source of light was brightest, is about 1.63; so that the primitive wave-front-theory gives 1.63 for the ratio of the speeds in air and CS_2, and the corpuscle-theory gives $(1.63)^{-1}$.

Foucault and Fizeau, be it remembered, had done the experiment with water. It happens that for water the derivative $dv/d\lambda$ is much smaller, and the group-speed therefore much closer to the wave-speed, than for carbon bisulphide. Also their experiments, though performed by the same method as Michelson was later to adopt and adapt, were less accurate than his. But if they had performed the Michelson experiment in 1850, the result would have been astounding. For Arago had asked, in effect: is it the speed of the wave-fronts in the wave-theory, or the speed of the corpuscles in the corpuscular theory, which agrees with the measured speed of a piece of light? Arago had said: "These experiments . . . will permit no further hesitation as between the rival theories. They will settle *mathematically* (I employ this word on purpose) they will settle mathematically one of the greatest and most disputed questions of natural philosophy." He had proposed a question to Nature, and had written down two and only two answers. Everyone thought that Nature must reply by ratifying one of the

---

[9] When the segments returned from the second to the first mirror they found that the latter had revolved a little further beyond the orientation which it had when they left it, so that it reflected them onward not quite along the path to the source of light, but along another path inclined to that one at an angle twice as great as that through which it had revolved. Michelson measured the angle, and knowing the rate of revolution of the revolving mirror he then knew how long the light had taken to go from it to the stationary mirror and back.

answers. Foucault and Fizeau reported that she had replied: *the former*. But they had not heard distinctly; for her actual response was: *neither*.

Michelson's experiment however came after the idea of group-velocity as distinguished from wave-velocity had been invented and established. The refractive index of carbon bisulphide varies with wave-length. On determining the wave-speed or phase-speed v from the refractive index (by the equation $N = c/v$) and then the derivative $dv/d\lambda$, it is found[10] that in the region of the visible spectrum, the term $\lambda(dv/d\lambda)$ amounts to about seven per cent of the term v, on the right-hand side of equation (20)—that is, the group-speed should be some seven per cent lower than the wave-speed in carbon bisulphide. In air, however, group-speed and phase-speed are sensibly the same. The ratio of the group-speeds in air and CS_2 falls close to Michelson's value.[11]

Coming as it did, therefore, the Michelson experiment merely showed that those who had subtilized the Huyghens' theory by introducing sine-waves had incidentally invented something able to move with the measured speed of a light-flash, though nothing of the sort had been available in the original form. Had it come earlier—well, there is no way of knowing what would have been inferred; but people might have come to think that after all a wavefront-theory or a corpuscle-theory of light may have some use and value, even though the speeds assigned to the waves or the corpuscles do not agree with those actually measured. Such an attitude of mind would be rather advantageous, today. As a corollary for the present I submit: in picturing a jet of free negative electricity as a beam of waves or a stream of corpuscles, we should not be too confident that either the speed of the waves or the speed of the corpuscles is the speed with which a segment dissected from the jet would move from place to place, until someone succeeds in making actual measurement of this last. Fundamental theory has something to say on this point, which we will presently consider.

---

[10] I take all the numerical values in this section from a review of Michelson's work by J. Willard Gibbs (*Am. Jour. Sci.* **31,** pp. 62–64; 1886) which so far as I know is the latest critical discussion of the data.

[11] The problem is more complex than I have intimated, not only because Michelson observed light covering a very wide range of wave-lengths so that v and $dv/d\lambda$ both extend over wide ranges of values, but also because different parts of a wave-front are reflected from different parts of the mirror at different moments, and therefore from *differently-inclined* parts. Quite a controversy went on during the eighteen-eighties in the pages of "Nature" as to what it was that Foucault had really measured. Rayleigh at first (*Nature* **24,** p. 382; 1882) thought it was g; then changed his mind, (**25,** p. 52; 1882) and decided it was v^2/g; then was convinced by Schuster (**33,** pp. 439–440; 1886) that it was really $v^2/2(v - g)$. J. W. Gibbs then took a hand (**33,** p. 582; 1886) and contended that after all it was really g. The controversy seems to have rested there. It may be added that Michelson's data eliminate v^2/g, but do not quite discriminate between g and Schuster's expression.

GROUP-SPEED AND CORPUSCLE-SPEED

Thus far I have said that if we wish to use wave-theory and corpuscle-theory alternatively, we must make the momentum of the corpuscle equal to the quotient of the constant h by the wave-length of the waves; but I have said nothing about the energy of the corpuscle.

Let us adopt the universal assumption—based on a multitude of experiments, for instance those on the photoelectric effect—that the energy E of a corpuscle of light is equal to the product of its frequency ν by the same universal constant h; and let us extend it to the other kinds of corpuscles which we may associate with other kinds of waves, and *vice versa*.

Then the complete description of the particles associated with waves of wave-length λ is as follows:

$$p = h/\lambda, \qquad E = h\nu = hv/\lambda. \tag{21}$$

Here, as before, v stands for the phase-speed of the waves (not the particles).

Returning to the formula (20) for the group-speed, we now can write it thus:

$$\begin{aligned} g &= v - \lambda(dv/d\lambda) = \nu\lambda - \lambda d(\nu\lambda)/d\lambda \\ &= -\lambda^2(d\nu/d\lambda) = -(\lambda^2/h)(dE/d\lambda). \end{aligned} \tag{22}$$

Suppose next that the energy and the momentum of the corpuscles in question are related to each other and to their speed in the well-known fashion of ponderable bodies, to which it is known that electrons conform. Thus for sufficiently low speeds, the relations are practically those of the "classical" mechanics:

$$p = m_0u, \qquad E = \tfrac{1}{2}m_0u^2, \text{ whence } E = p^2/2m_0. \tag{23}$$

Here m_0 stands for the constant mass, u for the speed of the corpuscles (*not* the waves).

The energy of the corpuscles is a function of the momentum only, and continuing to develop the formula (22) for the group-speed, we find:

$$\begin{aligned} g &= (-\lambda^2/h)(dE/dp)(dp/d\lambda) = dE/dp \\ &= p\sqrt{1 - \beta^2}/m_0 = u. \end{aligned} \tag{24}$$

The group-speed of the waves is equal to the speed of the corpuscles.

The same conclusion follows if we use the relativistic definitions for the energy and the momentum of a particle,

$$E = m_0c^2/\sqrt{1 - \beta^2}, \quad p = m_0\beta c/\sqrt{1 - \beta^2} \qquad (\beta = u/c),$$
$$E = c\sqrt{m_0^2c^2 + p^2}$$

as the reader may test for himself.

Summarizing: if the corpuscles associated with the waves have the properties of ordinary material bodies—*if*, let us say, for short, *the corpuscles are material particles, their speed is equal to the group-speed of the waves.*

This is a very happy and agreeable result. It compensates very largely for our having been forced to concede that if we want both waves and corpuscles, the wave-speed and the corpuscle-speed must be different. The wave-theory has supplied another velocity which is equal to that of the corpuscles. Moreover it is precisely the velocity with which we should expect an isolated segment of a wave train to move from place to place. If someone were to cut a piece out of an electron-jet and measure the time it took to traverse a known distance, the speed which he would deduce from his data would probably agree both with the corpuscle-speed and with the group-speed, and disagree with the wave-speed. It would be interesting to try this out.

In the equations (23) I have taken account only of the kinetic energy of the corpuscles; in the equations (25), only of their kinetic energy and of the "rest" energy associated with their mass. But the explanations of refraction by the two theories will no longer be concordant, unless the potential energy also is admitted. Let us denote the potential energy of a corpuscle by U; and, since as yet these theories have been verified only for negative electricity, let us immediately write eV for U, e standing for the charge of an electron and V for the electrostatic potential in the region where it is. For the total energy of the corpuscle, then, we have instead of (25) the relativistic expression,

$$E = m_0c^2/\sqrt{1 - \beta^2} + U = m_0c^2/\sqrt{1 - \beta^2} + eV, \qquad (26)$$

which for small values of the corpuscle-speed u $(= \beta c)$ reduces to the classical expression,

$$E = \tfrac{1}{2}m_0u^2 + U = \tfrac{1}{2}m_0u^2 + eV. \qquad (27)$$

In an earlier section we interpreted the refraction of an electron beam passing from vacuum into metal by thinking of the metal and the vacuum as being two regions in which different values of electrostatic potential prevail, the potential thus changing sharply from one value to the other at the surface which bounds the solid. Now when the beam considered as a stream of corpuscular electrons passes across such a surface, the energy of each electron as expressed by (26) or (27) remains the same, though the proportion which is kinetic energy is changed; and therefore the frequency E/h of the equivalent wave-train remains the same. If then we keep the assumption that the wave-length of the waves is equal to h divided by the momentum of the particles, we have the following value for the ratio between the wave-speeds v' and v on the two sides of the surface:

$$v/v' = \nu\lambda/\nu\lambda' = \left(\frac{E}{h}\,\frac{h}{p}\right)\Big/\left(\frac{E}{h}\,\frac{h}{p'}\right) = p'/p, \tag{28}$$

and the speed of the waves varies inversely as the momentum of the corpuscles, which is just what is required in order that we may hold both the theories simultaneously.

But how about the theorem that corpuscle-speed is equal to group-speed? Returning to the equations (25), we see that the introduction of the potential energy has altered the relation between energy and momentum; we now have:

$$E = c\sqrt{m_0c^2 + p^2} + eV. \tag{29}$$

But so long as we are comparing different electron-streams in the same medium (vacuum, for instance), the potential energy is the same for all and does not depend on the momentum; and differentiating E with respect to p to obtain the value of the group-speed g, we get:

$$g = dE/dp = \frac{c^2p}{E - eV} = \frac{c^2m_0u/\sqrt{1-\beta^2}}{m_0c^2/\sqrt{1-\beta^2}} = u, \tag{29}$$

and thus group-speed and corpuscle-speed are equal, as before.

I will write down the expression of the phase-speed, although for the physicist it is of minor importance, not being measurable—a fact which exempts us, temporarily at least, from pondering over the curious feature that it depends on the value of the potential energy of the corpuscles, and therefore (for electrons) on the value accepted for the electrostatic potential of the region where the wave-train is, even though in practice it is generally assumed that electrostatic potential may be measured from an *arbitrary* zero. The formula is this:

$$\begin{aligned} v = E/p &= \frac{m_0c^2/\sqrt{1-\beta^2} + U}{m_0u/\sqrt{1-\beta^2}} \\ &= c^2/u + U/p, \end{aligned} \tag{30}$$

and if we put the potential energy of the corpuscles equal to zero, we find the phase-speed varying inversely as the corpuscle-speed,[12] and greater than the speed of light.

---

[12] There is a paradox here which, as I can testify from personal experience, is a dangerous source of confusion. The formula $v = c^2/u$ sounds like an approximation to the formula $v = const/p$ which I have given as the requisite relation between

STATIONARY WAVES AND OSCILLATING PARTICLES

We have tried out, separately and in tandem, two alternative ways of interpreting a beam of radiation advancing through space; first as a stream of corpuscles, then as a train of waves. We will now try out two alternative ways of interpreting radiation enclosed in a box; first as a system of stationary waves, then as a quantity of corpuscles rushing to and fro and bouncing from the walls. To simplify the case as much as possible, think only of motions parallel to one side of the box; or to make the pictures more graphic, think of a tube or pipe like those often used in experiments on sound, in which the waves travel along the axis.

Now it is well known that when a train of sound-waves is sent through a tube, or generated by vibrations somewhere in the tube, it is partially reflected from the far end, then again partially reflected from the near end, and so on over and over again; the overlapping wave trains passing to and fro interfere with one another; and when the wave-length is related in a certain way to the length of the tube, the overlapping wave trains form a *stationary wave-pattern* of alternating loops and nodes—the tube is said to be in resonance. If the two ends of the tube are alike (both open, or both closed) so that reflection takes place in the same way as both, the waves which admit of resonance are those of which the half-wave-length or an integer number of half-wave-lengths fits exactly into the tube; denoting by d the length of the tube, these wave-lengths are given by the formula:

$$n\left(\frac{\lambda}{2}\right) = d, \qquad n = 1, 2, 3, \ldots \tag{41}$$

This equation defines what may be called the *characteristic wave-lengths* of the tube. The tube distinguishes these, or the wave trains possessing these wave-lengths, from all the others.

Suppose on the other hand we had particles rushing back and forth along the axis of the tube, and rebounding without loss of energy whenever they struck either wall. Denote by u the speed of a particle; it takes a time-interval $2d/u$ to describe a complete round-trip with two rebounds, and one might say crudely that it has a frequency $u/2d$. I say "crudely" because the corpuscle is not moving with a sinusoidal motion, like a pendulum-bob; its speed does not vary as a sine-function

wave-speed and momentum. However the two relate to entirely different situations. The first is a comparison between wave-speeds and corpuscle-speeds for different beams in the same medium. The second is a comparison between wave-speeds and corpuscle-momenta for the same beam in different media. The resemblance between the two is accidental and misleading.

I am indebted to Professors C. H. Eckart and E. C. Kemble for elucidation of this point.

of time, but retains the same value throughout except for the change of direction; if we were to apply a Fourier analysis to this motion, we should find not only the frequency $u/2d$, but all of its overtones. Let us think however only of this fundamental frequency. Now it is evident that there is nothing, in our ordinary conceptions of particles rushing back and forth and rebounding from walls, to distinguish any value of speed or frequency above any others. The phenomenon of resonance sets certain wave-lengths apart from others, but there is nothing to correspond to resonance in this latter case, and set certain speeds apart from others.

But instead of sound, think of some kind of radiation which we have interpreting both as corpuscles and as waves—light, for example. Light enclosed between parallel reflecting walls forms stationary waves,[13] provided that its wave-length is related to the distance d between the walls by the equation (41), which I rewrite:

$$\lambda = 2d/n, \qquad n = 1, 2, 3 \ldots \tag{42}$$

The parallel reflecting walls, or the limitation which they set upon the space accessible to the light, thus single out certain characteristic wave-lengths and distinguish them from all others. How interpret this fact by corpuscle-theory?

Well, we have been associating waves of wave-length λ with corpuscles of momentum $p = h/\lambda$; let us continue to do so. The reflecting walls, then, single out certain characteristic values of momentum given by this equation, derived straight from (42):

$$p = nh/2d, \tag{43}$$

which I proceed to rewrite thus,

$$2d \cdot p = nh \qquad n = 1, 2, 3 \ldots \tag{44}$$

These values of momentum, I have said, are set apart from all the rest. If waves and corpuscles are interchangeable as bases for a theory of light, then the feature of wave-motion known for short as "resonance" obliges us to make that supposition. But in what way, and to what extent, are they set apart? According to modern quantum-theory, they are actually the *only* possible values. A particle describing a cyclic motion of this character, in which it moves a fixed distance with a fixed momentum and then moves the same distance backward with the same momentum reversed and so forth ad infinitum, is constrained by something in the order of nature to have one or

---

[13] Interference patterns are essentially of this type, though usually they are formed between mirrors oblique to one another.

another of the "permitted" momenta defined by equations (43) and (44).

Examining equation (44), one sees how this definition of the permitted momenta may be stated. The quantity on the left of (44) is the product of the momentum of the particle, by the distance which it traverses each time it performs its cycle.[14] This product must be equal to an integer multiple of the Planck constant h.

Now the quantum-theory of the atom developed fifteen years ago by Bohr, Sommerfeld and W. Wilson—the first and greatest of the forward steps in the contemporary conquest of the problem of atomic structure—was based on the assumption that an electron performing a cyclic motion must perform it in such a way, that its momentum conforms to a condition of which equation (44) is but a special case. This is the condition always written thus:

$$\int pdq = nh, \qquad n = 1, 2, 3 \ldots \tag{45}$$

If the electron is oscillating to and fro in a straight line through a position of equilibrium, q stands for its distance from that position and p for its momentum, and the integral is taken once around a complete oscillation. It is evident that (44) is the special form of this equation for the case in which the force acting on the electron is vanishingly small until it hits the wall and then suddenly becomes enormous. If the electron is revolving in an orbit in two or three dimensions, there are two or three equations like (45) all postulated at once; but I shall not take up such more complicated cases.

Summarizing the outcome of this section in a phrase: *if we associate waves of wave-length λ with corpuscles of momentum h/λ, and stationary waves in an enclosure with corpuscles flying back and forth between its walls, then the condition that the waves must fulfil to form a stationary system is equivalent to the quantum-condition imposed upon the corpuscles.*

This is an illustration of wave-mechanics. How extraordinarily fruitful and valuable such comparisons have proved in the hands of Louis de Broglie, of Schroedinger, Bose, Fermi and Sommerfeld—to name only a few—I have shown in part, in earlier issues of this journal. Here it must suffice to say that Schroedinger developed the principle into a form suitable for predicting the stationary states of atoms; Bose constructed out of it a competent theory of radiation in thermal equili-

---

[14] It travels a distance d in the forward sense with a momentum p, and then an equal distance in the backward or negative sense with a momentum of equal amount but reversed sign, so that the total product of distance by momentum is

$$pd + (-p)(-d) = 2dp.$$

brium, considered as a gas of which the atoms are corpuscles of light; while Fermi, Dirac and Sommerfeld between them used it to make a powerful theory of the free negative electricity in metals, conceiving this alternatively as a gas of which the atoms are electrons, and a system of stationary waves enclosed within the surface of the metal as in a box with reflecting walls.

DIFFRACTION OF WAVES AND DIFFRACTION OF CORPUSCLES

The effect of a diffraction-grating upon a beam of light projected against it has always been considered the most striking evidence that light is of the nature of waves and not of corpuscles. Indeed it is considered to suffice in itself to prove the corpuscle-theory untenable. With any common understanding of the term *corpuscle-theory*, this statement is correct; but we had better put it in the softer form, that the effect of a diffraction-grating on a beam of light proves that if we adopt a corpuscular theory we must endow the corpuscles with some very strange property which nobody ever thought that particles could possess, and which may even seem to be in contradiction with their nature. We had better put the statement in this milder way, because it now is known that in spite of all the evidence for individual electrons, a beam of negative electricity is affected by a grating in much the same way as a beam of light.

Take then almost the simplest conceivable case of diffraction; a plane-parallel beam of light falling perpendicularly on a wall containing many equally-spaced parallel slits, and a part of the light passing through the slits to a screen infinitely far away. On this infinitely-distant screen—which may in practice be brought up to a convenient nearness, by means of a lens—one sees a peculiar pattern of light and shade. I single out one particular feature of this pattern: the fact that there are maxima of illumination along certain lines parallel to the slits. One of these, for instance, is straight ahead from the slits, along the direction of the incident beam prolonged; another is off to one side, in a direction making a certain angle (say ϕ) with that of the incident beam; another is equally far off to the other side. These two last-named, the *first-order maxima*, are those we shall consider; it will be enough to speak of one.

By the wave-theory, a first-order maximum is explained as follows. Each of the slits is the source of a secondary wave train of spherical wave crests, stimulated by the primary wave train, and having the same frequency and wave-length. Consider any two adjacent slits. Secondary wave crests start from the two at the same moment. At any point equally distant from the slits, they arrive simultaneously, and reinforce each other; this is the explanation of the central bright fringe. At any point not quite equally distant from the slits, they do not arrive quite simultaneously, and the reinforcement is impaired. But at a point which is further from one slit than from the other by just the wave-length λ, the wave crest arriving from the latter meets the next previous crest from the former, and the reinforcement is restored. The first-order maximum is located at these points.

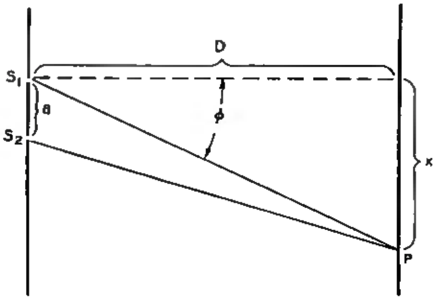


Fig. 4.

From Fig. 4 one sees[15] that when the screen is very far away, the points distant from the slit S_1 by one wave-length more than they are distant from S_2 are situated in the direction inclined at ϕ to the straight-ahead direction, the angle ϕ being given approximately by the formula

$$\sin \phi = \lambda/a, \tag{46}$$

where a stands for the distance between the slits. When the screen is infinitely far away, the formula is exact. (I must admit that it is somewhat disingenuous to simplify the problem by solving only the special case in which the screen is infinitely far away, for the general case opposes much more serious difficulties to the corpuscle-theory; but this is the special case of greatest physical importance, and one has to make a beginning somewhere.)

We have now explained the presence of a first-order maximum in the pattern of light and shade on the screen, though it cannot be said that we have "verified" formula (46), for that formula serves as the practical definition of wave-length; wave-lengths are measured by

---

[15] From the figure we see that for d_1 and d_2, the distances from S_1 and S_2 to the point P on the screen, we have:

$$d_1^2 = D^2 + x^2, \qquad d_2^2 = D^2 + (x-a)^2, \qquad d_1 = D \sec \phi, \qquad x = D \tan \phi$$

and hence

$$(d_1 - d_2)(d_1 + d_2) = 2ax - a^2.$$

When D, x, d_1 and d_2 all become infinite together, the second factor on the left becomes equal to $2D \sec \varphi$ and the second term on the right may be neglected.

measuring the angle ϕ and using equation (46). Let us now try the corpuscle-theory on the problem.

Putting as heretofore the value h/λ for the momentum of the corpuscles, translate (46) into the language of the alternative theory; one gets:

$$\sin \phi = h/ap. \tag{47}$$

In words: a corpuscle of momentum p, passing through any slit, is particularly likely to bend around through an angle ϕ of which the sine depends in a certain way on its momentum and on the distance to the next slit.

Which is to say: the likelihood that a corpuscle entering a slit will bend its course through a certain angle depends on the presence of other slits in the same wall, and on the distance between these slits.

But the reader will inquire: how does the corpuscle entering one of the slits know that the other slits are there? If all the other slits were suddenly stopped up, the first-order maximum would vanish, the likelihood that the corpuscle would turn in the direction given by (47) would fall to zero; but how could it know that they had been stopped up?

Well! this is precisely the strange and extravagant property with which we are forced to endow the corpuscles, if we want to use the particle-theory to explain diffraction. It must be supposed that when passing through a slit, a particle of light knows whether there are other slits and, if so, how they are spaced. It must be supposed that an X-ray particle striking an atom in a crystal knows that there are other atoms in a regular array, and knows moreover just the pattern and the scale of that array. It must be supposed that electrons enjoy a like omniscience. Or to express it in more technical language; the probability that a corpuscle of light, of electricity or of matter shall be deflected through a given angle when it strikes an atom or passes through a slit must be supposed to depend on the arrangement of the other atoms or the other slits in the vicinity. This idea is not easy to accept; but it must be accepted, if one is to build up a complete corpuscular theory of any of these entities.

But if one accepts it, one finds that the stipulation (47) turns out to be another example of the general quantum-condition of which, in (44), we have already met one instance. For write it thus:

$$ap \sin \phi = ap_t = nh, \qquad n = 1, 2, 3 \ldots, \tag{48}$$

the factor n being now introduced to take account of the maxima of second, third, and higher order which also occur on the screen, though

I refrained from mentioning them earlier. I have used the symbol p_t for the quantity $p \sin \phi$, for this, as one sees immediately, is the tangential component of momentum which the corpuscle acquires at the deflection, not having had any before. The wall containing the slits, or the row of atoms if we consider instead the diffraction of X-rays by a crystal, receives an equal momentum in the opposite sense. We may therefore say that diffraction occurs in such a way, that the regularly-spaced series of slits or atoms receives a momentum p_t given by the formula:

$$ap_t = nh. \tag{49}$$

But now what is the product ap_t? It is the product of the momentum of the row of atoms or slits by the distance a between any adjacent two; it is therefore the integral $\int pdq$ of the general principle (45), evaluated for the range of integration a. Now the general principle is supposed to apply when the range of integration covers a complete cycle of a periodic motion. There is nothing obviously periodic about a steady sidewise sliding of a row of atoms with a constant momentum. But in a sense, there is after all something periodic. For if the row of equally-spaced atoms (or slits) extends to infinity in both directions, then when it has moved sidewise through the distance a each atom lies exactly in the former place of another atom, and the original arrangement is to all appearances restored. The steady onward motion of the regular array is also a cyclic departure and return to a periodically-restored arrangement; and the maxima of the diffraction-pattern are determined by applying the quantum-condition to this cyclic motion.

The reader may ask: how about the component of momentum in the direction at right angles to the grating? Without precisely answering that question, I will end the article by applying the corpuscular theory to a case in which all the components of momentum are duly taken into account: diffraction of X-rays or of electrons by a three-dimensional crystal.

Suppose an "ideal" crystal extending infinitely far in all directions. It is composed of similar and similarly-oriented "atom-groups"—I will use the language and the symbols of the eighteenth article of this series—arranged upon a "space-lattice," of which the three spacings shall be denoted by a, a', a''. If we start with one atom-group A, then along one direction from it there is an infinite sequence of such groups at distances a, $2a$, $3a$, . . . and also at distances $(-a)$, $(-2a)$, $(-3a)$, . . . in the opposite sense. Call that the x-direction. Then along another direction through A, say the y-direction, there is an infinite sequence of groups at distances a', $2a'$, $3a'$, . . . and $(-a')$, $(-2a')$, etc.; and along a third or z-direction through A, there is an infinite sequence of atom-groups spaced at intervals a''.

Now think of the atom-groups as hard particles, and the corpuscle of light or of electricity (the "X-ray quantum" or the electron) as a hard particle which rushes into the lattice, hits one of the atom-groups—A, say—and bounces off. Denote by ϕ, ϕ', ϕ'' the angles which its original direction of motion makes with the x, y, z directions respectively; by θ, θ', θ'' the angles which its final direction of motion makes with these three. Before the deflection, the corpuscle has a momentum of magnitude p, parallel to its original direction of flight; afterward it has a momentum of the same magnitude, but parallel to its final direction of flight. At the deflection, then, it loses—that is, it communicates to the lattice—a momentum of which the three components along x, y, z have the values:

$$p(\cos\theta - \cos\phi); \qquad p(\cos\theta' - \cos\phi'); \qquad p(\cos\theta'' - \cos\phi'').$$

Now if, following the foregoing procedure, we equate the first of these to some integer multiple of h/a, the second to some integer multiple of h/a', and the third to some integer multiple of h/a'', and then translate momentum of corpuscles into wave-length of waves by the now-familiar formula $p = h/\lambda$, we get:

$$\begin{aligned} a(\cos\theta - \cos\phi) &= n\lambda, \\ a'(\cos\theta' - \cos\phi') &= n'\lambda, \\ a''(\cos\theta'' - \cos\phi'') &= n''\lambda, \end{aligned} \tag{50}$$

where n, n', n'' stand for any three integers. Now these are the equations (numbered 3, 4, 5 in the eighteenth article) to which conform the "Laue beams," which is to say, the directions in which electrons and light are actually diffracted by crystals.

Perhaps I should close with two or three admonitions. To make the wave-theory and the corpuscle-theory equivalent for a few simple cases is of course not at all the same as making them equivalent universally. Also, the examples in this article are not always so elementary as they may seem. The first involved two distinct media with a sharp boundary between; and discontinuity is always less agreeable than continuity to the mathematician. The last but one involved a non-sinusoidal vibration, which is much more complex than a sinusoidal one. Moreover, the concepts of light-waves and quanta are not nearly so beautifully welded together as those of electricity-waves and electrons. Nevertheless these illustrations may help to weaken the idea that there is no way out of the present situation but to abandon either waves or corpuscles; for decidedly, there is a way.

Wave Propagation Over Continuously Loaded Fine Wires

By M. K. ZINN

The paper contains the results of a theoretical investigation of wave propagation along a pair of wires that are "loaded" by enclosing each wire in a continuous sheath of magnetic material. The results of greatest practical interest are certain approximate formulas that are sufficiently simple to be adapted to engineering design studies, while having a high degree of precision for all practical dimensions and frequencies.

THE purpose of this investigation is to define the character of wave transmission along a pair of wires each of which is loaded with a continuous sheath of magnetic material. Exact expressions for the propagation constants are developed from the general theory that applies to such a system. Also, simple approximate formulas are given for the sizes of wires that are generally used in paper-insulated cables.

Wave Propagation Along a Pair of Wires with Magnetic Sheaths

For the benefit of those who are not interested in following the theoretical work in detail, a general sketch of the method and a summary of the mathematical results will be given first, together with a discussion of some numerical examples. Details of the theoretical work have been placed in the Appendices.

The analysis here given follows closely the methods developed by John R. Carson [1] in a solution of the transmission of periodic currents along a system of coaxial cylinders. The analysis for the case where the outgoing and return conductors are coaxial is applied, with only small modifications, to the case where the two conductors are parallel and not coaxial. This application of the theory ignores the "proximity effect." [2] That is to say, it assumes that the electric and magnetic forces within each conductor are functions only of the distance from its axis and of the coordinate in the direction of propagation, which is strictly true where the cylindrical conductors are coaxial.

---

[1] "Transmission Characteristics of the Submarine Cable," John R. Carson and J. J. Gilbert, *Journal of the Franklin Institute*, December 1921.

[2] This is the usual method of dealing with problems involving balanced parallel conductors. The alternating-current resistance of the system may be expressed as the product of the a.c. resistance, assuming a concentric return, and a "proximity effect correction factor," which takes into account the influence of the parallel return conductor. The "proximity effect" is in general negligible at voice frequencies for conductors of sufficiently small cross-section, such as those of paper insulated cables. References: "Wave Propagation over Parallel Wires: The Proximity Effect," John R. Carson, *Phil. Mag.*, April 1921, and "Wave Propagation over Parallel Tubular Conductors," Sallie Pero Mead. *Bell System Technical Journal*, April 1925.

The physical system contemplated is shown in Fig. 1. The outgoing and return systems of conductors, each comprising a cylindrical wire with insulated cylindrical sheath, are assumed to be identical in all respects. For the sake of generality, it is assumed that the magnetic sheaths may be insulated from the wires, as shown. The interesting practical case where wire and sheath are contiguous, forming a bi-metallic conductor, then appears as the limiting case of infinitesimally thin insulation.

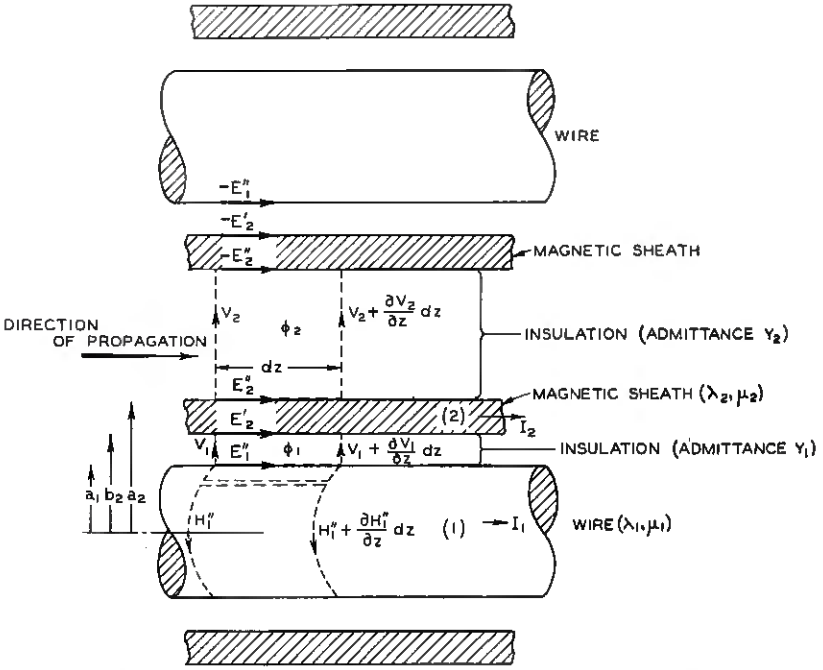


Fig. 1—Illustrating various quantities involved in the analysis.

The problem consists in finding a solution for the propagation constant of the system from Maxwell's equations. If the magnetic sheaths are in contact with the wires, the propagation constant is given in the usual form, $\Gamma = \sqrt{Y_2 Z}$, where Y_2 is the admittance across the insulation between the sheaths and Z is the series impedance of the system. The admittance is, in general, either a known, or an experimentally determined, quantity; so that for this case the theoretical problem resolves itself into that of finding the series impedance.

An important part of the investigation is, however, to determine what the effect would be of introducing insulation between the wire and its sheath. In this more general system, shown by the sketch, the solution for the propagation constant has two values, because two layers of insulation are involved, and cannot be expressed in the usual form. It is found, however, that it can be expressed in terms of the propagation constant for the elementary case where wire and sheath are in contact by introducing two other known propagation constants that determine transmission along the separate pairs of conductors in the system. The expression for the propagation constant, when given in this form, shows directly the effect of insulating the wires from their sheaths.

It is necessary first to define certain impedances. Let I_1 be the total current in one of the wires and I_2 the total current in its sheath. The tangential electric forces in the surfaces of wire and sheath are denoted by E_1'', E_2' and E_2'', as shown in Fig. 1. These electric forces are linear functions of the currents, as follows:

$$\begin{aligned} E_2'' &= Z_{21}''I_1 + Z_{22}''I_2, \\ E_2' &= Z_{21}'I_1 + Z_{22}'I_2, \\ E_1'' &= Z_{11}''I_1. \end{aligned} \qquad (1)$$

The impedances which appear in these equations as the coefficients of the currents are functions of the electrical constants and dimensions of the wires and sheaths. Their values are given in Appendix A.

Now let

γ = propagation constant determining transmission along the loaded wires if the wires and their sheaths were in contact $= \sqrt{Y_2 Z}$.

γ_{12} = propagation constant determining transmission along one wire with its sheath as the return, when the sheath is insulated from the wire $= \sqrt{Y_1 Z_{12}}$.

γ_{22} = propagation constant determining transmission along the two sheaths if the wires were removed $= \sqrt{Y_2 Z_{22}}$.

Then, from (1)

$$\left.\begin{aligned} Z_{12} &= \frac{E_2' - E_1''}{I_2} + X_1 = Z_{11}'' - Z_{21}' + Z_{22}' + X_1, && I_2 = -I_1 \\ Z_{22} &= \frac{2E_2''}{I_2} + X_2 = 2Z_{22}'' + X_2 && I_1 = 0 \\ Z &= \frac{2E_2''}{I_1 + I_2} + X_2 = 2\left[Z_{22}'' - \frac{Z_{22}'^2}{Z_{11}'' - Z_{21}' + Z_{22}'}\right] + X_2 && \end{aligned}\right\} \qquad (2)$$

In these equations,

$X_1 = i\omega L_{12}$ = reactance arising from the magnetic field between the outer surface of the wire and the inner surface of the sheath.

$X_2 = i\omega L_{22}$ = reactance arising from the magnetic field between the two sheaths.

The terms in brackets in the equation for Z give the "internal impedance" of one of the loaded wires for the elementary case where wire and sheath are in contact, and X_2 is the additional reactance that arises from the magnetic field outside the wires.

With the elementary propagation constants, γ, γ_{12} and γ_{22} so defined, it is found that the propagation constant, Γ, of the general system can be expressed as follows:

$$2\Gamma^2 = \gamma_{12}^2 + \gamma_{22}^2 \pm \sqrt{(\gamma_{12}^2 + \gamma_{22}^2)^2 - 4\gamma^2\gamma_{12}^2}. \tag{3}$$

It is convenient also to express the two solutions for Γ in the form of series:

$$\Gamma_1^2 = \gamma^2 \frac{\gamma_{12}^2}{\gamma_{12}^2 + \gamma_{22}^2} + \gamma^4 \frac{\gamma_{12}^4}{(\gamma_{12}^2 + \gamma_{22}^2)^3} + 2\gamma^6 \frac{\gamma_{12}^6}{(\gamma_{12}^2 + \gamma_{22}^2)^5} + \cdots,$$
$$\Gamma_2^2 = \gamma_{12}^2 + \gamma_{22}^2 - \Gamma_1^2. \tag{4}$$

The solutions in the series form show the effect of introducing insulation between the wire and sheath. For, if $\left|\frac{4\gamma^2\gamma_{12}^2}{(\gamma_{12}^2 + \gamma_{22}^2)^2}\right|$ is small compared to unity, as it would be in a continuously loaded wire with a thin magnetic sheath of high resistance, then, to a first order of approximation, the principal propagation constant Γ_1 is less than γ, the propagation constant that determines transmission when wire and sheath are in contact, by the factor

$$\sqrt{\frac{1}{1 + \frac{\gamma_{22}^2}{\gamma_{12}^2}}}.$$

The other propagation constant, Γ_2, is, in this case, very large compared to Γ_1 and plays no appreciable part in defining the character of transmission except at points very near to the terminals of the system. For practical purposes, the system may be considered to have only one significant mode of propagation.

CASE OF A WIRE WITH CONTIGUOUS SHEATH

The Internal Impedance

The practical case where the magnetic sheath and the wire are contiguous, forming a bi-metallic conductor, is of special interest.

In this case, the propagation constant is uniquely determined from a knowledge of the admittance between the loaded wires and of their series impedance. The "internal impedance" of the loaded wires comprises the larger part of this impedance. For the purpose of engineering design work, it is convenient to have at hand approximate formulas for the "internal impedance."

The exact expression for the impedance is given by the last of equations (2). When the magnetic sheath is thin, as compared to the radius of the copper wire, certain approximations can be made. These are explained in Appendix A. The result is the following formula for the "internal impedance":

$$\frac{Z_i}{2} = \frac{1 - \omega^2 F + i\omega G}{\frac{1}{R} + i\omega H}, \tag{5}$$

where $F = \pi\mu_2 b\left[\frac{4}{3}\pi\lambda_2\lambda_1\mu_2 t^3 + \mu_1 b\left(\frac{\lambda_1}{2R_2} - \frac{\lambda_2}{R_1}\log\frac{a}{b}\right)\right]$,

$$G = \frac{\mu_2}{R_2} + \frac{\mu_1}{R_1} + \pi b^2 L_2(\lambda_1 - \lambda_2),$$

$$H = 2\pi^2\lambda_2\mu_2 a t^2\left[\frac{2}{3}t(\lambda_2 - \lambda_1) + \lambda_1 b\right] + \frac{\mu_1}{2R_1R_2},$$

$R = \dfrac{R_1R_2}{R_1 + R_2}$ = d.-c. resistance of one of the pair of bi-metallic conductors,

$R_1 = \dfrac{1}{\pi\lambda_1 b^2}$ = d.-c. resistance of the inner part of the conductor (the wire),

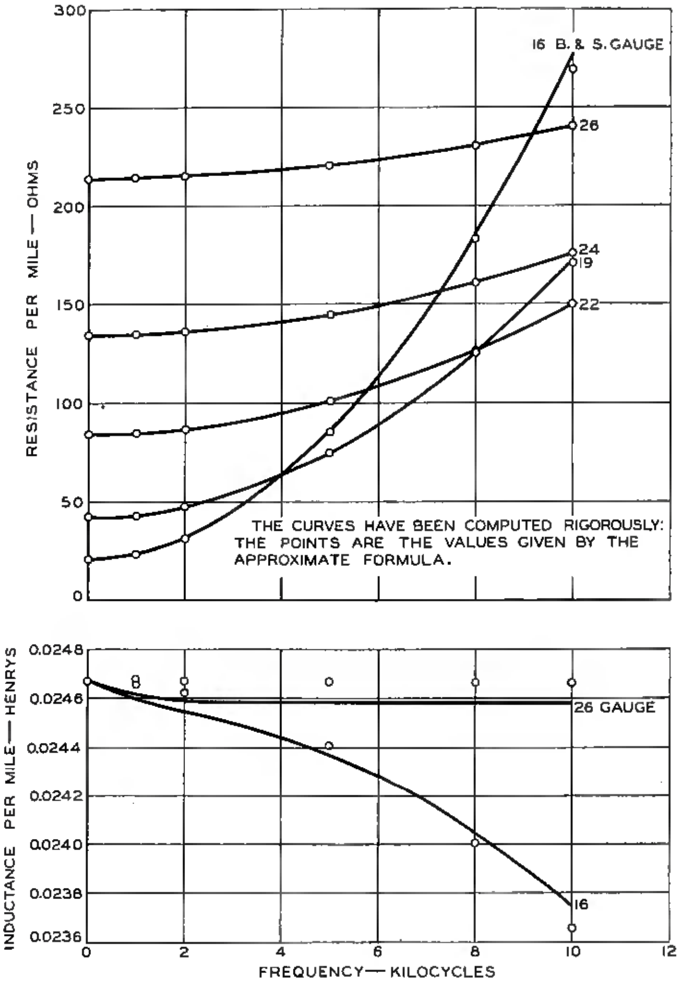
$R_2 = \dfrac{1}{\pi\lambda_2(a^2 - b^2)}$ = d.-c. resistance of the outer part of the conductor (the sheath),

$L_2 = 2\mu_2 \log\dfrac{a}{b}$ = low-frequency inductance contributed by the sheath,

b = radius of the wire,
a = outside radius of the sheath,
$t = a - b$ = thickness of the sheath,
λ_1, μ_1 = conductivity and permeability of the wire,
λ_2, μ_2 = conductivity and permeability of the sheath,
$\omega = 2\pi$ times the frequency,
$i = \sqrt{-1}$.

The total series "loop" impedance of the pair of loaded conductors per centimeter is $Z = Z_i + X_2$.[3]

For the purpose of indicating the degree of precision of the approximate formula, data are given in Fig. 2 on the internal resistance and inductance of various copper wires coated with loading material to



THE CURVES HAVE BEEN COMPUTED RIGOROUSLY: THE POINTS ARE THE VALUES GIVEN BY THE APPROXIMATE FORMULA.

Fig. 2—Internal impedance of wires of various sizes with continuous loading of approximately 25 millihenrys per wire mile (for very small currents, i.e., hysteresis losses not included).

---

[3] All quantities are expressed in the electromagnetic c.g.s. system of units. To obtain the result in ohms per loop mile, multiply by 160,934 (10^{-9}). In the case of cable circuits, X_2 ($= i\omega L_{22}$) is an experimentally determined quantity, L_{22} having a value of about .001 henry per mile.

such a depth as to give an internal inductance of about .025 henry per wire mile. The magnetic material in the sheath has been assumed to have a permeability of 3,000 and a conductivity of .77 x 10^{-4} in e.m.u. (resistivity 13 microhm-centimeters, in practical units). The data shown by the solid lines are exact while the points give the results obtained by means of the approximate formula (5). A comparison of results is tabulated below for the largest wire (16 B. & S. gauge), where the errors of the approximate formula are greatest.

Frequency—Kilocycles	Internal Resistance and Inductance (of One Wire) Ohms and Millihenrys per Mile					
	Exact		Approximate		Errors	
	Res.	Ind.	Res.	Ind.	Res.	Ind.
0	21.065	24.77	21.065	24.77	—	—
2	31.674	24.56	31.63	24.63	−.14%	+.29%
5	86.795	24.37	86.18	24.41	−.71	+.16
8	186.65	24.05	183.6	24.01	−1.63	−.17
10	276.04	23.75	269.4	23.66	−2.41	−.39

The errors are roughly proportional to the quantity, $t\sqrt{\omega\mu_2\lambda_2}$. For a loading material having, say, one-quarter the permeability and the same conductivity, the errors would be about twice as large, therefore, if the inductance and the wire size remain the same.

Hysteresis Loss

The real part of the internal impedance given by (2) or (5) is the effective internal resistance of the bi-metallic wire, taking into account the heat losses that arise from the electric current, namely, d.-c. resistance, eddy current loss and "skin effect loss." The formulas do not take into account hysteresis loss, which is a magnetic phenomenon as distinguished from these electric phenomena. The determination of hysteresis loss rests upon experimental data. If the energy loss due to hysteresis in the magnetic material per unit volume per cycle is h (ergs), then the resistance increment due to hysteresis is

$$R_h = \frac{\omega}{I^2}\int_b^a hrdr. \tag{6}$$

For the low values of magnetic force that obtain in telephony, it is found that $h \doteq \eta B^3$, where η is the hysteresis coefficient and B the induction density. Therefore

$$R_h = \frac{\eta\omega\mu^3}{I^2}\int_b^a H^3rdr. \tag{7}$$

Since the magnetic coating is thin, and the "demagnetizing," or "screening," effect of eddy currents small, it may be assumed that $H = 2I/r$. (It will not exceed that value, at least.) Using this approximate value for H, the resistance increment due to hysteresis is

$$R_h \doteq 8\eta\omega\mu^3 I \frac{a-b}{ab} \doteq 2\eta\mu\omega B_a L_2, \tag{8}$$

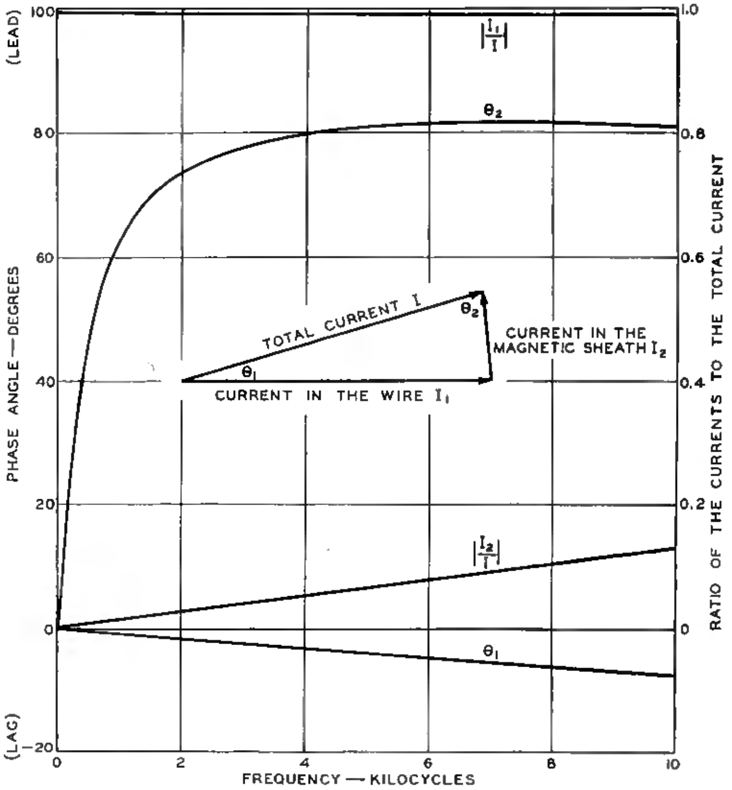


Fig. 3—Illustrating the fractions of the total current that are carried by the copper wire and by the magnetic sheath (19 gauge (B. & S.) wire with continuous loading of 25 millihenrys per mile).

where B_a is the induction density at the outside boundary of the sheath.

The Distribution of the Current in Wire and Sheath

It is a matter of interest to know how much of the current is carried by the magnetic sheath and how the current is distributed over the cross-section of the wire and sheath at various frequencies. The solution of this problem is not an essential part of the investigation, but it helps in understanding what takes place in the bi-metallic conductor.

The ratio of the currents in wire and sheath to the total current, as computed from (1), is plotted in Fig. 3. It will be noted that the fraction of the total current carried by the sheath becomes greater as the frequency increases. But the fraction carried by the copper nevertheless remains very nearly unity at all frequencies. This behavior is explained by the curves representing the phase angles involved. These show, of course, that at very low frequencies the

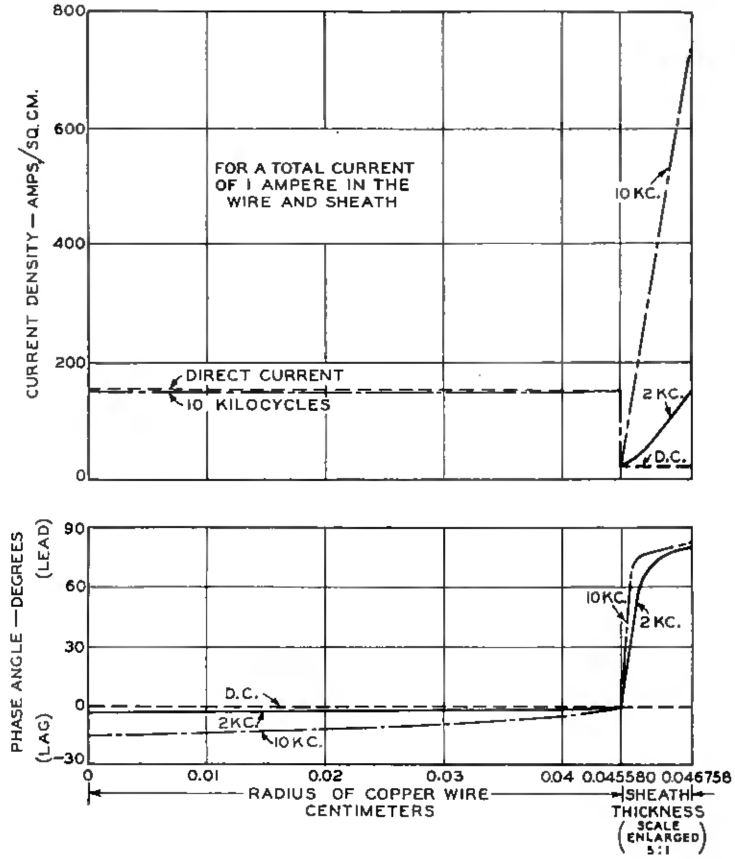


Fig. 4—Illustrating the current density throughout the cross-section of a wire loaded with a continuous magnetic sheath—for direct current and 2 and 10 kilocycle alternating currents. (Same 19 B. & S. gauge wire as that of Fig. 3.)

copper current and the sheath current are nearly in phase, but with increasing frequency, the copper current lags behind the sheath current, until at high frequencies the two currents approach a quadrature phase relation.

It may be said that at high frequencies the current in the loading material is practically all "wattless" current, in the sense that it contributes very little to the energy delivered to any receiving device connected to the line, but it dissipates energy, of course. At 10 kilocycles, for the 19-gauge loaded wire, the current carried by the magnetic sheath contributes only 2 per cent of the useful current (see Fig. 3); yet 75 per cent of the energy loss occurs in the sheath (see Fig. 2).

The difference in phase between the component currents in wire and sheath is explained by the consideration that the reactance of a given filament of current is proportional to the magnetic flux external to it. In the copper, therefore, the elementary current paths have a small resistance, but a large reactance, due to the fact that nearly all the magnetic flux is in the loading material. Near the outer surface of the loading material, on the other hand, the current paths have less internal reactance, but the resistance is large.

This brings the discussion to Fig. 4, which shows how the amplitude and phase of the current varies over the cross-section of the bi-metallic conductor for direct current and for 2 and 10 kilocycle alternating currents. For the 19-gauge loaded wire, illustrated, the "skin effect" in the copper is seen to be very small, the alternating current distribution being practically uniform, as for direct current. At the boundary between the copper and the magnetic material, the current amplitude suffers a discontinuity, but the phase is continuous. The discontinuity in the current amplitude conforms to the law that the component of electric force along the conductor must be continuous at a boundary, which requires that the ratio between the current amplitudes on the two sides of the boundary must equal the ratio of the conductivities of the two materials. The current density distribution over the cross-section of the magnetic sheath is uniform for direct current, of course, but for alternating currents, the density increases and the phase advances abruptly toward the outer surface of the sheath.

APPENDIX A

The impedances [4] which appear in equation (1) in the body of the paper as the coefficients of the currents are given by:

---

[4] See above noted paper (reference 1) for the development of these formulas.

$$
\begin{aligned}
Z_{21}'' &= \frac{2i\omega\mu_2}{x_2}\,\frac{U_2 - 1}{U_2'}, \\
Z_{22}'' &= \frac{2i\omega\mu_2}{x_2}\,\frac{U_2}{U_2'}, \\
Z_{21}' &= \frac{2i\omega\mu_2}{x_2}\,\frac{1 - \dfrac{a_2}{b_2}V_2'}{U_2'}, \\
Z_{22}' &= \frac{2i\omega\mu_2}{x_2}\,\frac{1}{U_2'}, \\
Z_{11}'' &= \frac{2i\omega\mu_1}{x_1}\,\frac{1}{U_1'} = \frac{2i\omega\mu_1}{x_1}\,\frac{J_0(x_1)}{J_0'(x_1)}
\end{aligned}
\tag{9}
$$

(Note that $Z_{22}' = Z_{22}'' - Z_{21}''$),

where

$$
\begin{aligned}
U_j &= -\,y_j[J_0(x_j)K_0'(y_j) - J_0'(y_j)K_0(x_j)], \\
V_j &= -\,y_j[J_0(y_j)K_0(x_j) - J_0(x_j)K_0(y_j)], \\
U_j' &= -\,y_j[J_0'(x_j)K_0'(y_j) - J_0'(y_j)K_0'(x_j)], \\
V_j' &= -\,y_j[J_0(y_j)K_0'(x_j) - J_0'(x_j)K_0(y_j)].
\end{aligned}
\tag{10}
$$

J_0 and K_0 are Bessel functions of zero order of the first and second kind, respectively, and J_0' and K_0' are their derivatives with respect to the arguments, which are given by

$$
\begin{aligned}
x_j &= a_j i\sqrt{4\pi i\omega\mu_j\lambda_j}, \\
y_j &= b_j i\sqrt{4\pi i\omega\mu_j\lambda_j},
\end{aligned}
\tag{11}
$$

where $\omega = 2\pi$ times the frequency, $i = \sqrt{-1}$, a_j and b_j are the outer and inner radii respectively of conductor j, and μ_j and λ_j are its permeability and conductivity. Quantities with the subscript 1 refer to the wire and those with the subscript 2 refer to the sheath. All quantities are expressed in the electromagnetic c.g.s. system of units.

Writing Maxwell's Law, curl $E = -\dfrac{d\Phi}{dt}$, around the contours indicated by the dotted rectangles in Fig. 1 gives

$$
E_2' - E_1'' - \frac{\partial V_1}{\partial z} = \frac{d\Phi_1}{dt}, \tag{12}
$$

$$
-\,2E_2'' - \frac{\partial V_2}{\partial z} = \frac{d\Phi_2}{dt}, \tag{13}
$$

where V_1, V_2 are the potential differences between the surfaces of the conductors, as shown, and Φ_1, Φ_2 are the normal values of the magnetic flux that cuts the surfaces bounded by the contours. The term $-2E_2''$ results from the symmetry of the system, which imposes the condition that the electric and magnetic forces at corresponding points in the outgoing and return conductors are equal and oppositely directed. Also, it is unnecessary to write a third equation for the field between the other wire and its sheath, because this equation would be the same as (12). Therefore, the transmission is characterized by only two modes of propagation.

Since all the variables are propagated at the same rate, and since sinusoidal currents are being considered, $\partial/\partial z$ may be replaced by $-\Gamma$ and d/dt by $i\omega$. Then

$$E_2' - E_1'' + \Gamma V_1 = X_1 I_1, \tag{14}$$

$$-2E_2'' + \Gamma V_2 = X_2(I_1 + I_2), \tag{15}$$

where Γ is the propagation constant and

$X_1 = i\omega L_{12} =$ reactance arising from the magnetic field between the outer surface of the wire and the inner surface of the sheath.

$X_2 = i\omega L_{22} =$ reactance arising from the magnetic field between the two sheaths.

The potential differences V_1, V_2 can be expressed in terms of the currents by writing Maxwell's Law, curl $H = 4\pi I$, around contours in the outside surfaces of wire and sheath. (Such a contour for the wire is indicated by dotted lines in the sketch.) This gives

$$2\pi a_1 \frac{\partial H_1''}{\partial z} = -4\pi V_1 Y_1, \tag{16}$$

$$2\pi a_2 \frac{\partial H_2''}{\partial z} = -4\pi V_2 Y_2. \tag{17}$$

where

$Y_1 =$ admittance across the insulation between wire and sheath.

$Y_2 =$ admittance across the insulation between the two sheaths.

Since $H_1'' = \dfrac{2I_1}{a_1}$ and $H_2'' = \dfrac{2(I_1 + I_2)}{a_2}$,

$$\Gamma I_1 = V_1 Y_1, \tag{18}$$

$$\Gamma(I_1 + I_2) = V_2 Y_2. \tag{19}$$

Substituting (18) and (19) in (14) and (15), respectively, gives

$$\left(X_1 - \frac{\Gamma^2}{Y_1}\right) I_1 = E_2' - E_1'', \tag{20}$$

$$\left(X_2 - \frac{\Gamma^2}{Y_2}\right) (I_1 + I_2) = -\,2E_2'', \tag{21}$$

and substituting (1) in (20) and (21) gives the two equations of the currents. In order that they shall be consistent, the determinant of the coefficients must vanish. Therefore

$$\begin{vmatrix} X_1 - \dfrac{\Gamma^2}{Y_1} - Z_{21}' + Z_{11}'' & -\,Z_{22}' \\[2ex] X_2 - \dfrac{\Gamma^2}{Y_2} + 2Z_{21}'' & X_2 - \dfrac{\Gamma^2}{Y_2} + 2Z_{22}'' \end{vmatrix} = 0. \tag{22}$$

The roots of this equation give the required solutions for the propagation constant. First, however, it is convenient to introduce two known propagation constants. Let

γ_{12} = propagation constant determining transmission along one wire with its sheath as the return $= \sqrt{Y_1 Z_{12}}$.

γ_{22} = propagation constant determining transmission along the two sheaths if the wires were removed $= \sqrt{Y_2 Z_{22}}$.

Then, from (1), (20) and (21),

$$\begin{aligned} Z_{12} &= Z_{11}'' - Z_{21}' + Z_{22}' + X_1, \\ Z_{22} &= 2Z_{22}'' + X_2, \end{aligned} \tag{23}$$

substituting (23) in (22) and rearranging,

$$\begin{vmatrix} \dfrac{\gamma_{12}^2 - \Gamma^2}{Y_1} & -\,Z_{22}' \\[2ex] -\,2Z_{22}' & \dfrac{\gamma_{22}^2 - \Gamma^2}{Y_2} \end{vmatrix} = 0. \tag{24}$$

Expanding

$$\Gamma^4 - \Gamma^2(\gamma_{22}^2 + \gamma_{12}^2) + \gamma_{12}^2\gamma_{22}^2 - 2Z_{22}'^2 Y_1 Y_2 = 0. \tag{25}$$

The remaining impedance can be eliminated by introducing γ, the propagation constant that would characterize transmission if the wires were in contact with the sheaths. (In order not to disturb the dimensions, it may be imagined that the insulation between wire and sheath be replaced by an infinitely conducting material, which, however, is assumed to conduct no current axially. Then $E_2' - E_1'' = X_1 I_1$.) To find γ, make Y_1 infinite and solve (25). Then

$$\gamma^2 = Y_2 Z,$$

and

$$Z = Z_{22} - \frac{2Z_{22}'^2}{Z_{12}}. \tag{26}$$

Therefore

$$2Z_{22}'^2 Y_1 Y_2 = \gamma_{22}^2 \gamma_{12}^2 - \gamma^2 \gamma_{12}^2. \tag{27}$$

Finally, substituting (27) in (25) and solving the resulting equation gives the two solutions for the propagation constant,

$$2\Gamma^2 = \gamma_{12}^2 + \gamma_{22}^2 \pm \sqrt{(\gamma_{12}^2 + \gamma_{22}^2)^2 - 4\gamma^2\gamma_{12}^2}. \tag{28}$$

The arbitrary constants remain to be determined. The currents are, in general,

$$\begin{aligned} I_1 &= A_{11}\epsilon^{-\Gamma_1 z} + A_{12}\epsilon^{-\Gamma_2 z} + B_{11}\epsilon^{\Gamma_1 z} + B_{12}\epsilon^{\Gamma_2 z}, \\ I_2 &= A_{21}\epsilon^{-\Gamma_1 z} + A_{22}\epsilon^{-\Gamma_2 z} + B_{21}\epsilon^{\Gamma_1 z} + B_{22}\epsilon^{\Gamma_2 z}. \end{aligned} \tag{29}$$

The condition of principal practical interest is that of a long cable with connection made to the two wires and with the sheaths left free at the sending end. For this case, the conditions are

$$\begin{aligned} &(1)\ \text{At } z = 0, \quad I_1 = I_0 \text{ and } I_2 = 0, \\ &(2)\ \text{At } z = \infty, \quad I_1 = 0 \text{ and } I_2 = 0, \end{aligned}$$

where I_0 is the current delivered to the cable pair at the sending end. From the second condition,

$$B_{11} = B_{12} = B_{21} = B_{22} = 0. \tag{30}$$

From the first condition,

$$\begin{aligned} A_{11} + A_{12} &= I_0, \\ A_{21} + A_{22} &= 0. \end{aligned} \tag{31}$$

But these constants must satisfy, for each of the two values of Γ, the equations of the currents, whose coefficients are given in the determinant (22). Therefore

$$\begin{aligned} A_{21} &= K_1 A_{11}, \\ A_{22} &= K_2 A_{12}, \end{aligned} \tag{32}$$

where

$$\begin{aligned} K_1 &= \frac{Z_{12} - Z_{22}' - \dfrac{\Gamma_1^2}{Y_1}}{Z_{22}'} = \frac{Z_{22} - 2Z_{22}' - \dfrac{\Gamma_1^2}{Y_2}}{- Z_{22} + \dfrac{\Gamma_1^2}{Y_2}}, \\ K_2 &= \frac{Z_{12} - Z_{22}' - \dfrac{\Gamma_2^2}{Y_1}}{Z_{22}'} = \frac{Z_{22} - 2Z_{22}' - \dfrac{\Gamma_2^2}{Y_2}}{- Z_{22} + \dfrac{\Gamma_2^2}{Y_2}}. \end{aligned} \tag{33}$$

Substituting (32) in (31) and solving

$$\begin{gathered} A_{11} = I_0 \frac{K_2}{K_2 - K_1}, \qquad A_{12} = - I_0 \frac{K_1}{K_2 - K_1}, \\ A_{21} = I_0 \frac{K_1 K_2}{K_2 - K_1} = - A_{22}. \end{gathered} \tag{34}$$

Finally, the currents are given by

$$\begin{aligned} I_1 &= \frac{I_0}{K_2 - K_1} \left[K_2 \epsilon^{-\Gamma_1 z} - K_1 \epsilon^{-\Gamma_2 z} \right], \\ I_2 &= \frac{I_0 K_1 K_2}{K_2 - K_1} \left[\epsilon^{-\Gamma_1 z} - \epsilon^{-\Gamma_2 z} \right]. \end{aligned} \tag{35}$$

This completes the analysis for the more general system where the magnetic sheaths are insulated from the wires. For the special case where wire and sheath are contiguous, γ_{12}^2 is infinite and (28) shows that $\Gamma_1 = \gamma$ and $\Gamma_2 = \infty$. The transmission is, therefore, defined by only one mode of propagation. The series impedance of the system is, from (23) and (26),

$$Z = 2 \left[Z_{22}'' - \frac{Z_{22}'^2}{Z_{11}'' - Z_{21}' + Z_{22}'} \right] + X_2, \tag{36}$$

where the terms in brackets give the internal impedance of one of the loaded wires, and X_2 is the reactance that arises from the magnetic field between them. The internal impedance can be obtained also by finding $\dfrac{2E_2''}{I_1 + I_2}$ directly from the last two of equations (1), of course.

The constant K_2 becomes -1 and the total current, $I = I_1 + I_2$, is propagated in accordance with $I = I_0\epsilon^{-\gamma z}$, where $\gamma = \sqrt{ZY_2}$.

The constant K_1, which is the ratio of the current in the sheath to that in the wire, is of interest. It becomes

$$\frac{I_2}{I_1} = K_1 = \frac{Z_{12} - Z_{22}'}{Z_{22}'} = \frac{Z_{11}'' - Z_{21}'}{Z_{22}'}. \tag{37}$$

The approximate formulas for the case where wire and sheath are contiguous are derived as follows: The arguments, x_2 and y_2, of the Bessel functions differ by only a small amount when the magnetic sheath is thin. This situation is favorable to an advantageous use of Taylor's series. $J_0(x_2)$, for example, can be expressed in terms of $J_0(y_2)$, its derivatives and the difference of the arguments in a Taylor series as follows:

$$J_0(x) = J_0(y) + \tau J_0'(y) + \frac{\tau^2}{2!}J_0''(y) + \frac{\tau^3}{6!}J_0'''(y) + \cdots, \tag{38}$$

where $\tau = x - y$ (x_2, y_2 being written simply, x, y, here, for convenience). Furthermore, Bessel functions are subject to recurrence formulas,[5] which enable us to express each of the derivatives occurring in the series in terms of the function of zero order, its first derivative and the argument. Therefore, by applying the recurrence formulas to the Taylor series, we find functions U and V (see Appendix B) such that

$$J_0(x) = UJ_0(y) + VJ_0'(y), \tag{39}$$

$$K_0(x) = UK_0(y) + VK_0'(y) \tag{40}$$

(U_2, V_2 being also written now, U, V). Differentiating (39) and (40) with respect to τ,

$$J_0'(x) = U'J_0(y) + V'J_0'(y), \tag{41}$$

$$K_0'(x) = U'K_0(y) + V'K_0'(y), \tag{42}$$

where $U' = \frac{\partial U}{\partial \tau}$, $V' = \frac{\partial V}{\partial \tau}$.

---

[5] The two recurrence formulas required are:

$$J_n'(z) = \frac{n}{z}J(z) - J_{n+1}(z),$$

$$J_n'(z) = J_{n-1}(z) - \frac{n}{z}J_n(z).$$

The Bessel Functions of the second kind satisfy the same formulas.

If (39) to (42) be solved for U, V, U', V', it can be verified that the solutions are the definitions of these functions already given in equations (10).[6]

The exact formula for the internal impedance of a wire with contiguous sheath has been given in (36). In terms of the functions U and V, this formula becomes

$$\frac{Z_i}{2} = \frac{2i\omega\mu_2}{x_2} \cdot \frac{\sqrt{\dfrac{\lambda_2\mu_1}{\lambda_1\mu_2}}\,\dfrac{U_2}{U_1'} + V_2}{\sqrt{\dfrac{\lambda_2\mu_1}{\lambda_1\mu_2}}\,\dfrac{U_2'}{U_1'} + V_2'}. \tag{43}$$

By using the series for these functions and discarding all terms of degree higher than ω^2, the approximation given in the body of the paper (equation 5) may be obtained.

APPENDIX *B*

When the recurrence formulas are applied to the Taylor series, it is found that

$$U = 1 + \frac{\tau^2}{2} + \frac{\tau^3}{6y} + \frac{\tau^4}{24}\left(1 - \frac{3}{y^2}\right) - \frac{\tau^5}{120}\left(\frac{2}{y} - \frac{12}{y^3}\right) - \frac{\tau^6}{720}\left(1 + \frac{10}{y^2} + \frac{60}{y^4}\right) + \cdots, \tag{44}$$

$$V = \tau - \frac{\tau^2}{2y} - \frac{\tau^3}{6}\left(1 - \frac{2}{y^2}\right) + \frac{\tau^4}{24}\left(\frac{2}{y} - \frac{6}{y^3}\right) + \frac{\tau^5}{120}\left(1 - \frac{7}{y^2} + \frac{24}{y^4}\right) - \frac{\tau^6}{720}\left(\frac{3}{y} - \frac{33}{y^3} + \frac{120}{y^5}\right) + \cdots. \tag{45}$$

These series converge for $\left|\dfrac{\tau}{y}\right| < 1$, which condition is satisfied by the sheath dimensions of any practical continuously loaded conductor.

A considerable number of the terms in the series for U and V are

---

[6] A relation that can be used to advantage at times is

$$U'V - UV' = -\frac{y}{x} = -\frac{b}{a}.$$

This relation corresponds to the similar one for the Bessel functions themselves, namely:

$$J_n'(z)K_n(z) - J_n(z)K_n'(z) = \frac{1}{z}.$$

parts of well-known series that define certain elementary functions. It can be verified readily that

$$U = \cos\tau + \frac{y^2}{2}\left[\log\left(1+\frac{\tau}{y}\right) - \frac{\tau}{y} + \frac{\tau^2}{2y^2}\right] - \frac{y^4}{12}\left[\log\left(1+\frac{\tau}{y}\right) - \frac{\tau}{y} + \frac{\tau^2}{2y^2} - \frac{\tau^3}{3y^3} + \frac{\tau^4}{4y^4}\right] + \cdots, \tag{46}$$

$$V = \sin\tau + y\left[\log\left(1+\frac{\tau}{y}\right) - \frac{\tau}{y}\right] + \frac{\tau^4}{12y} - \frac{7\tau^5}{120y^2} - \frac{\tau^6}{240y} + \frac{11\tau^6}{240y^3} + \cdots, \tag{47}$$

$$U' = -\sin\tau + \frac{y}{2}\,\frac{\left(\frac{\tau}{y}\right)^2}{1+\frac{\tau}{y}} + \frac{\partial}{\partial\tau}\text{(above remainder of (3))}, \tag{48}$$

$$V' = \cos\tau - \frac{\frac{\tau}{y}}{1+\frac{\tau}{y}} + \frac{\partial}{\partial\tau}\text{(above remainder of (4))}. \tag{49}$$

The series (46) to (49) possess a certain advantage for computing in that the quantities in brackets are real numbers. $\left(\text{Note that } \frac{\tau}{y} = \frac{a_2 - b_2}{b_2}.\right)$ They have been used also in obtaining the approximate formulas given in the body of the paper.

The quantities discussed above all pertain to the sheath. For finding U_1', involved in the last of the formulas (9), the series are not valid, of course. For this we have the well-known series,

$$\frac{1}{U_1'} = \frac{J_0(x_1)}{J_0'(x_1)} = \frac{1}{x_1}\left[-2 + \frac{x_1^2}{4} + \frac{x_1^4}{96} + \frac{x_1^6}{1536} + \frac{x_1^8}{23040} + \cdots\right] \tag{50}$$

—see, e.g., Gray, Mathews and McRoberts, "Bessel Functions," 2d edition, p. 170.

Theory of Vibration of the Larynx[1]

By R. L. WEGEL

The vibration in the larynx is caused by an automatic modulation by the vocal cords of the air stream from the lungs. Analytically the mechanism is the same, and physically, closely analogous to that of the vacuum tube oscillator. It depends principally on the resonance of the vocal cords, the modulation of air friction in the glottis by their motion and the attraction due to constriction of the air stream between them. When these forces exist in certain relative proportions and phases, sustained oscillation as in singing takes place. The whole mechanism may be represented analytically by force equations, from which conditions for accretion or subsidence of the vibration or for sustained oscillation may be easily deduced. These equations also show the analogy with other types of oscillating systems.

IT is customary in treating the theory of the voice to assume the glottis or space between the vocal cords to be a source of a steady stream of air with superimposed periodic impulses caused by the vibration of the vocal cords. The harmonic content of these impulses is modified by the "resonating" vocal cavities before being radiated into free air. It is the nature of this modification which receives most attention. The mechanism by which the vibration of the vocal cords is maintained has not been carefully studied.

The vocal cords are maintained in a state of sustained vibration by the proper balance between the various mechanical constants of the complete system, which thus act as a transformer of a part of the non-vibratory power derived from the air stream from the lungs into the vibratory power resulting in sound. It is a simple theory of this mechanism which is considered here.

The method used is to obtain the force equations, which describe the vibrations of the complete mechanical system, by means of the Lagrange equations, from expressions of the total instantaneous kinetic and potential energies, the instantaneous forces acting and rate of dissipation of energy. The resulting simultaneous equations relating to the displacements and velocities of the various parts are then studied to find the frequencies of free vibration and the relations which must obtain between the various mechanical parameters of the system in order that one of these frequencies be sustained. The method is an application of the theory of H. W. Nichols, published in *Physical Review*, August, 1917.

The theory is reduced to easily workable form by the introduction of simplifying approximations which will be described in the progress

---

[1] Presented before Acoustical Society of America, May 11, 1929.

of the discussion. The principal one of these is the neglecting of all reactions of second or higher order, thus leaving a set of linear differential equations.

STRUCTURE OF THE VOCAL TRACT

The vocal tract consists of three principal parts, the lungs and associated respiratory muscles for maintaining a flow of air, the larynx (see Fig. 1) for producing the periodic modulation and the upper vocal cavities, pharynx, mouth and nose for varying the relative harmonic content of sound originating in the larynx.

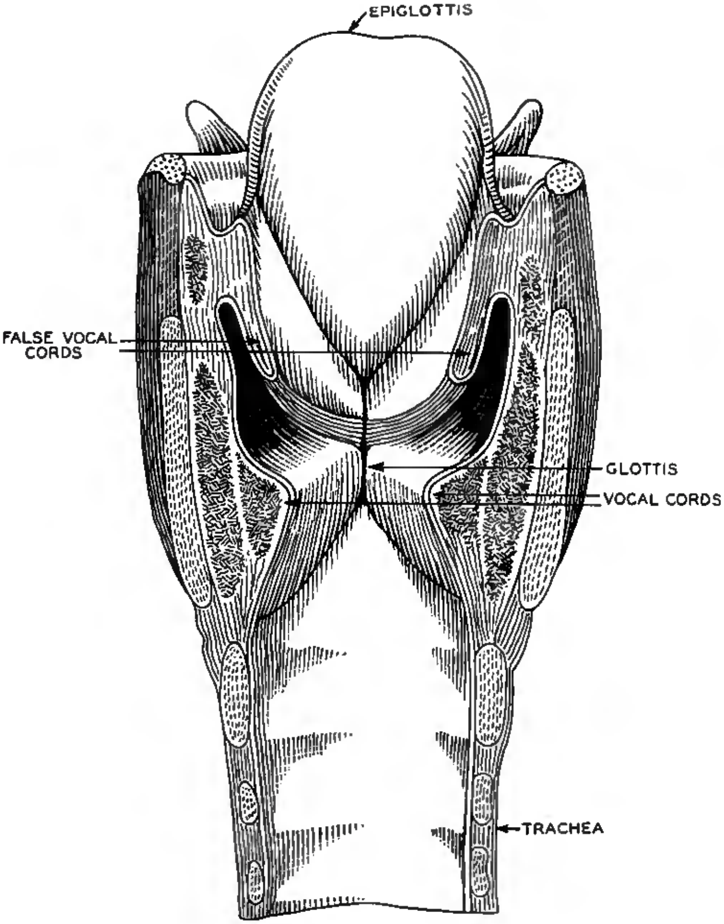


Fig. 1—Anterior-Posterior Section of the Larynx.

The capacity of the lungs in an adult man is capable of being varied from about two to five liters. The average in quiet breathing is about 2.6 liters. The average expiration of air in quiet breathing is about .5 liter. The rate of expiration of air in medium loud singing varies from 40 to 200 cm.^{3}/sec., the lower values obtaining for trained singers.

The larynx (see Fig. 1) consists of an irregularly shaped cartilaginous box at the top end of a tube, the trachea, about 12 cm. long by 2 cm. in diameter, leading from the lungs. The larynx contains the vocal cords, a pair of fibrous lips which in vibrating vary the width of the slit called the glottis, between them. The length of the glottis in the adult male averages about 1.8 cm. and in the female 1.2 cm. The width of the glottis varies widely with differing sounds. A few tenths of a millimeter may be considered representative. The tension and separation of the vocal cords are controlled by muscles.

The principal upper vocal cavities are the pharynx, a space just over the larynx, the mouth and the nasal cavities. The first and second may be varied in size and shape at will, but the effect on the last is controlled only by varying the communicating aperture between it and the pharynx.

EQUATIONS OF MOTION OF THE LARYNX

Fig. 2 shows a cross-section of a model which illustrates the essential details of the larynx in so far as it is necessary for this treatment. S_0 represents the area of the opening to the trachea. The vocal cords are represented by elastically hinged members of combined effective area S_2. By effective area is meant the area of aperture which displaces the same volume of air as the vocal cords when it moves the distance q_2 of the tips of the cords. This area is less than that of the vocal cords. The tips of the vocal cords are separated to form a gap, the glottis, of area S_1. A positive or up and outward displacement q_2 of the vocal cords increases S_1. It will be assumed that the air is not appreciably compressed in the neighborhood of the glottis, that is, any tendency to compression is relieved by flow into the trachea or pharynx.

The pressure in the lungs forces a steady current of air through the glottis. Let the velocity in the trachea of this steady flow be I_0 and in the glottis I_1. Small vibrations of the vocal cords superimpose additional small velocities, i_0 and i_1, in the trachea and glottis respectively. If the instantaneous velocity of the vocal cords be i_2 and it be assumed that they are constrained to move in synchronism

$$(I_0 + i_0)S_0 = (I_1 + i_1)S_1 + i_2S_2. \tag{1}$$

The above material is a description of a simple model of two degrees of freedom which simulates the principal characteristics from the standpoint of performance of the more complex larynx which has many degrees of freedom. It is this idealized model which will be considered in the subsequent treatment. Such points of performance of the actual larynx which may be due to the action of ignored and presumably subsidiary modes of motion will, of course, not be predicted by the theory. These are assumed to be of minor importance. The possible independence of motion of the two vocal cords will be considered later, however.

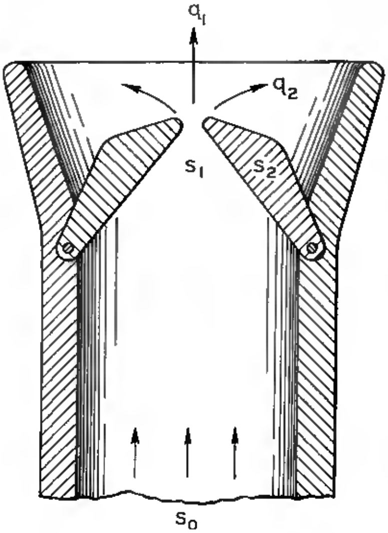


Fig. 2—Schematic Larynx Model.

The contraction of the air stream at the glottis introduces a relatively large concentrated kinetic energy in the air stream at this point similar to that at the mouth of a Helmholtz resonator. The inertia of a small plug of air between the vocal cords may then to a first approximation be treated as a mass L_1. A concentration of frictional resistance also occurs at this point due to viscosity and to turbulence. A positive displacement q_2 (outward) of the vocal cords causes an increase in the mass of the plug of air in the glottis and a change in the effective resistance, R, encountered by it. The inertia L_1' and resistance R of the glottis are therefore both functions of q_2, the displacement of the vocal cords from a mean position, and of the width of the glottis. If further Q_2 represent the average displacement of the vocal cords from an appropriately chosen reference position, L_2 their inertia coefficient and K_2 their effective stiffness, all measured at the tips, the total kinetic energy, T, and potential energy, V, of the larynx are

$$T = \tfrac{1}{2}L_1'(I_1 + i_1)^2 + \tfrac{1}{2}L_2 i_2^2, \tag{2}$$

$$V = \tfrac{1}{2}K_2(Q_2 + q_2)^2. \tag{3}$$

The Lagrange equation of forces for the nth coordinate of any system is

$$E_n = \frac{\partial}{\partial t}\frac{\partial T}{\partial i_n} - \frac{\partial T}{\partial q_n} + F_n + \frac{\partial V}{\partial q_n}, \tag{4}$$

in which F_n is a reaction due to friction. The force equations for the glottis and vocal cords therefore become

$$E_0 = F_1 + \frac{\partial}{\partial t}\frac{\partial}{\partial i_1}\left[\tfrac{1}{2}L_1'(I_1 + i_1)^2\right], \tag{5}$$

$$0 = L_2\frac{di_2}{dt} + F_2 + K_2(Q_2 + q_2) - \frac{(I_1 + i_1)^2}{2}\frac{\partial L_1'}{\partial q_2}. \tag{6}$$

NATURE OF THE "CONSTANTS" OF THE SYSTEM

It is quite safe to conclude that none of the coefficients (inertia, dissipation and stiffness) of the larynx are sensibly constant over the range of operation of the coordinates. Direct measurements are evidently impossible. It is conceivable that they may be arrived at indirectly by means of a comparison of experimental data, especially taken for the purpose, on voice curves and the results of dynamic analysis of the kind described here. The problem may also be studied by means of models. In order to solve equations 5 and 6 it is, however, necessary to evaluate the space and velocity derivatives.

A few simple experiments were performed on models for the sole purpose of determining the qualitative nature of variation of resistance of the glottis with displacement of the vocal cords. A diagram of the model used in the measurements is shown in Fig. 3. This consists of a brass tube, *a*, ¾″ in diameter, beveled off on the top at an angle of 45° with the axis, and two ⅛″ brass plates, *b*, fitted on these beveled surfaces so as to leave a slit, *S*, which was made adjustable in width. A cross-section of this model is shown in *c*. The bottom of the tube was attached to a large air chamber in which the pressure and velocity of air flow could be regulated and measured.

Three shapes of "glottis" were measured. The first had square corners, as shown on Fig. *3c*. The second, *3d*, was the same as *3c*, except that the corners of the lips were rounded. The third, Fig. 3*e*, had square corners as before, but the slit was about .1 mm. wider in the middle than at the ends.

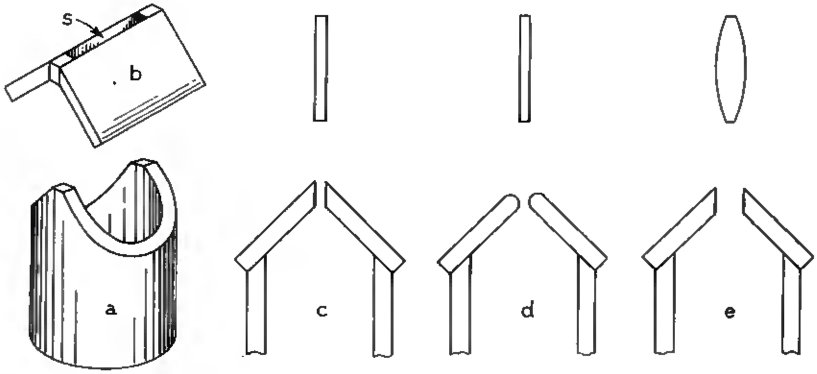


Fig. 3—Glottis Models.

The resistance R is given as the ratio of the product of pressure and slit area to the linear velocity of flow. Measurements were made in each case through a range of pressures such as to give fluxes through the slit through a range of 50 to 200 cm.³/sec. (Stanley and Sheldon values, see *Sci. Am.*, Dec. 1924) and through a range of slit width W of .01 to .10 cm. The data can be represented approximately in this range for the three slits by the following formulæ:

$$R = 3.6 \; I^2 W^{1.7} \times 10^{-6},$$
$$R = 6.1 \; I^2 W^{2.2} \times 10^{-6},$$
$$R = 800 \; I^2 W^{2.3} \times 10^{-6}.$$

In these expressions I is the velocity of flow of air through the slit. More careful data taken through a wider range of I and W would undoubtedly have given R in a power series.

These formulæ are taken to indicate that the resistance of the actual glottis increases faster than a linear function of I and W due to turbulence and may be represented as a single valued function of either displacement of the vocal cords q_2 (or glottis width) or of air velocity as expressed by a Taylor's series as follows:

$$R = R_0 + \frac{\partial_0 R}{\partial q_2} q_2 + \frac{\partial_0 R}{\partial \dot{i}_1} \dot{i}_1 + \tfrac{1}{2}\left(\frac{\partial_0{}^2 R}{\partial q_2{}^2} q_2{}^2 + \frac{2\partial_0{}^2 R}{\partial q_2 \partial \dot{i}_1} q_2 \dot{i}_1 + \frac{\partial_0{}^2 R}{\partial \dot{i}_1{}^2} \dot{i}_1{}^2 \right) + \text{etc.} \quad (7)$$

In this expression R_0 is the resistance measured in the reference position at which point the derivatives are taken, where i_1 and q_2 are zero. The experiment mentioned above determines the signs of the coefficients of q_2 and i_1 as positive. If the flow were purely laminar, i.e. due to viscosity only, the first would be negative and the second zero.

THE REACTION F_1

By definition, $E_0 = R_0 I_1$, where E_0 is the force of the lung pressure on the glottis and I_1 a corresponding linear velocity of flow of air. If a force F_1 slightly greater than E_0 act on the glottis and result in a velocity $I = I_1 + i_1$,

$$\frac{F_1}{I_1 + i_1} = R. \tag{8}$$

A combination of (7) and (8) constitutes an evaluation of F_1 for substitution in the force equation (5). To a first order approximation then:

$$F_1 = R_0 I_1 + \left(R_0 + I_1 \frac{\partial_0 R}{\partial i_1} \right) i_1 + I_1 \frac{\partial_0 R}{\partial q_2} q_2. \tag{9}$$

The coefficient of q_2 is dimensionally a stiffness and that of i_1 a resistance. In what follows they will be denoted by

$$F_1 = R_0 I_1 + R_1 i_1 + K_u q_2. \tag{10}$$

GLOTTIS MASS (L_1') REACTIONS

The kinetic energy of the air stream being proportional to the volume integral of the square of the velocity is largely concentrated in the glottis on account of the relatively high velocity at this point. On account of the irregularity in shape and turbulence in the stream it is impracticable to attempt an integration. If the velocity were so small that the turbulence were absent an approximate value of the air mass would be obtained by taking the mass of a cylinder of air having the length of the slit and a diameter equal to its width. This would make the mass L_1' proportional to W^2, or since the width is proportional to displacement of the vocal cords, to q_2^2. Owing, however, to turbulence and other non-linearities, the mass is probably more nearly described as a tongue of air issuing from the glottis, the inertia L_1' of which varies as some power function of the width and also of the velocity.[2]

---

[2] It has been found since experimentally that the mass reaction is very nearly that of a cylinder as described but reduced somewhat in diameter due to viscous or turbulent drag at the tips of the vocal cords.

It might be seen by carrying through an expression for this glottis mass involving a function of velocity similar to that for R of equation (7) that only a quantitative change in effective mass would result in the final equations and that no new type of reaction would be introduced. This demonstration is not included here. In order to save space in this qualitative treatment it is ignored. For small displacements q_2 from a reference position at which the velocity of the air is I_0, the glottis inertia may be represented by the direct function:

$$L_1' = L_1 + \frac{\partial_0 L_1}{\partial q_2} q_2 + \tfrac{1}{2} \frac{\partial_0^2 L_1}{\partial q_2^2} q_2^2 + \text{etc.,} \tag{11}$$

in which the coefficient of q_2 is obviously positive. The second term of the second member of (5) may now be evaluated by performing the differentiations as indicated. Neglecting second and higher order terms and denoting dq_2/dt by i_2 the reaction in question becomes

$$L_1 \frac{di_1}{dt} + I_1 \frac{\partial_0 L_1}{\partial q_2} i_2. \tag{12}$$

The glottis mass of air, therefore, introduces two kinds of reactions: a simple inertia and a reaction proportional to the velocity of the vocal cords. For simplicity of notation (12) will be written

$$L_1 \frac{di_1}{dt} + Gi_2. \tag{13}$$

This completes the evaluation of the terms (5), the force equation of the glottis, which may now be written

$$E_0 = R_0 I_1 + R_1 i_1 + K_u q_2 + \frac{L_1 di_1}{dt} + Gi_2. \tag{14}$$

FORCE EQUATION OF THE VOCAL CORDS

The force equation (6) of the vocal cords contains four terms. The first is the inertia reactance of the vibrating lips. The mass L_2 is the effective vibrating mass which, if multiplied by one-half the square of the velocity at the cord tip, gives the kinetic energy of their motion. If the distribution of the velocity in the vocal cords were known this might be found by integration. The second term F_2 in equation (6) represents the internal dissipation and is assumed proportional to the small velocity i_2. The third term is the elastic reaction which is proportional to displacement.

The fourth term is a "gyrostatic" term. This term may be written as follows:

$$-\frac{\partial T}{\partial q_2} = -\tfrac{1}{2}(I_1 + i_1)^2\left(\frac{\partial_0 L_1}{\partial q_2} + \frac{\partial_0{}^2 L_1}{\partial q_2{}^2}q_2 + \text{etc.}\right). \tag{15}$$

Again by neglecting second and higher order effects this reaction becomes

$$-\frac{I_1{}^2}{2}\frac{\partial_0 L_1}{\partial q_2} - I_1\frac{\partial_0 L_1}{\partial q_2}i_1 - \frac{I_1{}^2}{2}\frac{\partial_0{}^2 L_1}{\partial q_2{}^2}q_2. \tag{16}$$

It will be seen that the first term of this expression represents a static force tending, since it is negative, to draw the vocal cords together. This is the Bernoulli effect utilized in a venturi meter. This steady force is counterbalanced by an elastic reaction of the vocal cords with which it combines to determine an equilibrium position which obtains when the cords are not vibrating. This term may, therefore, be dropped from the final equations representing only superimposed motions.

The coefficient of i_1 is identical, except for a sign, with G of (13). It represents a force on the vocal cords due to a superimposed part of the Bernoulli effect caused by the small superimposed velocity i_1 in the glottis. The coefficient of q_2 is dimensionally a stiffness. This apparent stiffness is due to the nature of the air flow and is independent of any elastic members. It is negative if the second differential of glottis mass with respect to cord displacement is negative, positive when this coefficient is positive and vanishes when this coefficient is zero. It simply adds or subtracts in effect from the stiffness K_2 of the vocal cords. The first possibility is the more likely.[3] These terms may then be written for simplicity

$$-\frac{\partial T}{\partial q_2} = -F - Gi_1 - K_x q_2. \tag{17}$$

FORCE EQUATIONS OF THE LARYNX

The force equations of the glottis and vocal cords with constants thus evaluated are

$$E_0 = L_1\frac{di_1}{dt} + R_1 i_1 + Gi_2 + R_0 I_1 + K_u q_2, \tag{18}$$

$$0 = L_2\frac{di_2}{dt} + R_2 i_2 + K_2 q_2 + K_2 Q_2 - F - Gi_1 - K_x q_2. \tag{19}$$

As explained before, $E_0 = R_0 I_1$ and $F = K_2 Q_2$; so these cancel and are of no interest here. In the following it will be seen that the field

---

[3] This coefficient has since been found to be negative.

stiffness K_x is included in K_2 to simplify notation. This leaves (18) and (19) finally:

$$0 = L_1 \frac{di_1}{dt} + R_1 i_1 + G i_2 + K_u q_2, \tag{20}$$

$$0 = L_2 \frac{di_2}{dt} + R_2 i_2 + K_2 q_2 - G i_1. \tag{21}$$

It should be noted that these equations represent all first order internal reactions of the idealized model of the larynx. The series expansions have been carried out, to show to what approximations these equations hold. It should also be pointed out that the effects of mechanical hysteresis of the parts, which make the relative positions of the parts dependent on the previous history of their motion, have not been considered. A consideration of hysteresis complicates the theory considerably and is ignored for the same reason and with the same justification and limitations that it is ignored in the elementary treatment of electrical circuits containing coils with magnetic material and condensers with electrostatic hysteresis.

EXTERNAL REACTIONS OF THE TRACHEA AND VOCAL CAVITIES ON THE LARYNX

So far the modifying effect of the trachea and lungs, as well as the upper vocal cavities, on the motion have not been considered. Before using the equations it is necessary to evaluate these reactions and add them in their proper places.

Imagine a weightless piston fitted into the trachea just below the vocal cords such that the volume of air thus enclosed in the larynx is so small in comparison to that of the trachea and lungs that its compressibility may be neglected. If the vocal cords are held rigid and the plug or piston of air in the glottis is forced inward, a reaction in addition to the resistance and inertia of the glottis will be encountered due to the impeding effect of the trachea piston, which impedance is determined by the constants of the lower chambers. If a small force f_0 act on the trachea, causing a small velocity, i_0, and we assume linearity of response $f_0 = Z_0 i_0$ where Z_0 is a constant which may, due to a positive inertia reactance or a stiffness, contributed by air compression in the lungs, involve either a time derivative or integral of displacement. For the present consider it to be a generalized impedance operator. Due to the relative incompressibility of the air in the larynx, the volume displaced by the trachea piston is $i_0 S_0 = i_1 S_1$. Since also the instantaneous pressure inside the larynx is constant on all its walls, including the surface of the trachea piston $f_0/S_0 = f_1/S_1$. We then have

$$f_1 = Z_0 \frac{S_1^2}{S_0^2} i_1. \tag{22}$$

This reaction due to the trachea must be added to those of the glottis given in (20). In like manner if the effective area of the vocal cords is S_2 a reaction f_2 must be added to their force equation

$$f_2 = Z_0 \frac{S_2^2}{S_0^2} i_2. \tag{23}$$

Due to the steady component of air flow there is a static component of pressure tending to force the cords outward. This is counter to the static Bernoulli term and again, if second order effects of small quantities be neglected, serves only to alter the equilibrium position and may therefore be disregarded here.

When the glottis plug of air is displaced inward a force is exerted on the vocal cords tending to move them outward which is relieved to a certain extent by a yield of the trachea piston. This force on the vocal cords may be shown by reasoning similar to that above to be

$$Z_0 \frac{S_1 S_2}{S_0^2} i_1. \tag{24}$$

Since this part of the system is linear, the reaction between glottis and vocal cords through this channel is reciprocal so a force is exerted on the glottis when the vocal cords are displaced of

$$Z_0 \frac{S_1 S_2}{S_0^2} i_2. \tag{25}$$

It will be noticed that S_1 is a variable because of the variation in width of the glottis while vibrating. The effect of this variation in these terms is obviously second order since i_1 is small and will therefore be neglected.

The reactions of the upper cavities might be similarly added, but they are apparently relatively small and since they are at present not quantitatively known, are disregarded in the general equations because of the increased complexity. Generally, however, Z_0 may be thought of as representing the additive effects of both upper and lower chambers.

The complete force equations of the voice for small vibrations, taking into account all major external as well as internal reactions, may then be written:

$$0 = L_1\frac{di_1}{dt} + R_1 i_1 + G i_2 + K_u q_2 + \frac{Z_0 S_1^2}{S_0^2} i_1 + \frac{Z_0 S_1 S_2}{S_0^2} i_2. \tag{26}$$

$$0 = L_2\frac{di_2}{dt} + R_2 i_2 + K_2 q_2 - G i_1 + \frac{Z_0 S_2^2}{S_0^2} i_2 + \frac{Z_0 S_1 S_2}{S_0^2} i_1. \tag{27}$$

These equations may be put in a somewhat simpler form by virtue of the fact that they are linear differential equations with constant coefficients. In such a case the time differential may be replaced by an algebraic operator p such that $i = pq$, $di/dt = p^2 q$, where p is of the dimensions and nature of a frequency

$$0 = \left(p^2 L_1 + pR_1 + pZ_0\frac{S_1^2}{S_0^2}\right) q_1 + \left(pG + pZ_0\frac{S_1 S_2}{S_0^2} + K_u\right) q_2, \tag{28}$$

$$0 = \left(-pG + pZ_0\frac{S_1 S_2}{S_0^2}\right) q_1 + \left(p^2 L_2 + pR_2 + K_2 + pZ_0\frac{S_2^2}{S_0^2}\right) q_2. \tag{29}$$

The determinant of this system is (calling $pZ_0 = Y_0$)

$$D = \begin{vmatrix} \left(p^2 L_1 + pR_1 + Y_0\frac{S_1^2}{S_0^2}\right) & \left(pG + Y_0\frac{S_1 S_2}{S_0^2} + K_u\right) \\ \left(-pG + Y_0\frac{S_1 S_2}{S_0^2}\right) & \left(p^2 L_2 + pR_2 + K_2 + Y_0\frac{S_2^2}{S_0^2}\right) \end{vmatrix}. \tag{30}$$

This determinant represents the complete reactions of the larynx and the external effects of communicating air chambers.

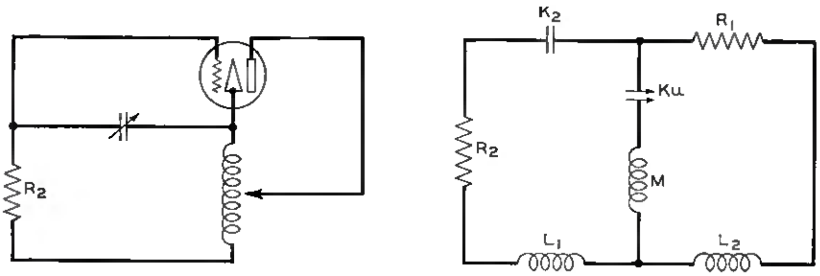
If the effects of the air chambers be disregarded the system is represented by placing $Y_0 = 0$, giving the simple form

$$D = \begin{vmatrix} (p^2 L_1 + pR_1) & (pG + K_u) \\ -pG & (p^2 L_2 + pR_2 + K_2) \end{vmatrix}. \tag{31}$$

NATURE OF THIS SYSTEM

The voice system represented by determinant (3) is very closely analogous to other vibrators, such as the microphone oscillator or door buzzer and the vacuum tube oscillator. The literature on the latter subject is now so extensive that the pointing out of the analogy should make the method of solution for sustained oscillation, as in singing, or for subsidence or accretion of the oscillation, as in speaking, clear to any one familiar with it.

Fig. 4*a* is a schematic diagram of a three-element vacuum tube oscillator circuit known as the "tuned grid" circuit. This is one of many kinds. The transformer coupling between the plate and grid circuit is represented by an auto-transformer. Fig. 4*b* represents the same circuit schematically but with circuit elements only. In this R_2 represents that part of the resistance of the coil which belongs to the grid circuit and any other associated dissipation, L_2 is the inductance of the coil as seen from the grid mesh and K_2 the reciprocal of the combined tuning capacity across the grid and that of the grid-filament. It is the electrical stiffness or elasticity of the grid mesh, in other words. R_1 is a plate-filament resistance ("a.c.") and L_1 that part of the coil in the plate circuit. M is the mutual inductance of the transformer which is not part of the mesh impedance of either plate or grid. The element K_u is the "uni-lateral mutual impedance" (G. A. Campbell, 1914) between the plate and grid meshes and is numerically equal to μK_2, where μ is the amplification constant of the tube. Other internal tube impedances are as usual neglected. The impedance determinant may be written directly from the circuit diagram Fig. 3*b*.



a *b*

Fig. 4—"Tuned Grid" Oscillator.

$$D = \begin{vmatrix} (p^2L_1 + pR_1) & (p^2M + K_u) \\ p^2M & (p^2L_2 + pR_1 + K_2) \end{vmatrix} . \qquad (32)$$

The quantities on the principal diagonal of this determinant, that is the first and last elements, are as usual in a circuit determinant the mesh impedances while the others are the mutuals. The principal features of the analogy may be seen by comparison of determinants (31) and (32). Except for the thus far undefined external or trachea impedance the mesh impedances are the same, from which it appears that the glottis is analogous with the plate-filament path in the vacuum tube and the vocal cords with the grid-filament path. The air velocity in the glottis I_1 corresponds to the plate current. In the vacuum tube this plate current is modulated by varying charge, q_2, on the grid. In the larynx the glottis air velocity is modulated by varying displacement, q_2, of the vocal cords. The charge on the plate (again neglecting internal capacities except the grid-filament) causes no effect on the grid mesh and in the larynx the position of any element of glottis air has no effect on the vocal cords. The unilateral mutual impedance, K_u, is the same in both.

The analogy breaks down at the point where the " feed back " part of the mechanisms is compared. The " feed back " is the bilateral part of the mutual impedance between the two meshes. In the vacuum tube circuit this is p^2M, the mutual of the transformer, while in the larynx it is pG, the "gyrostatic" mutual. The latter is a type of element which does not occur in electrical circuits, arising as it does from a variation of a mass or inductance with a displacement. Inductance, being a function purely of the geometry of a circuit, can only vary with mechanical displacement and not with electrical displacement or charge. The gyrostatic mutual is common in the mechanics of rotating bodies whence it derives its name. It is also the mutual in an electromagnetic telephone receiver or relay between the electrical circuit and the armature or diaphragm.

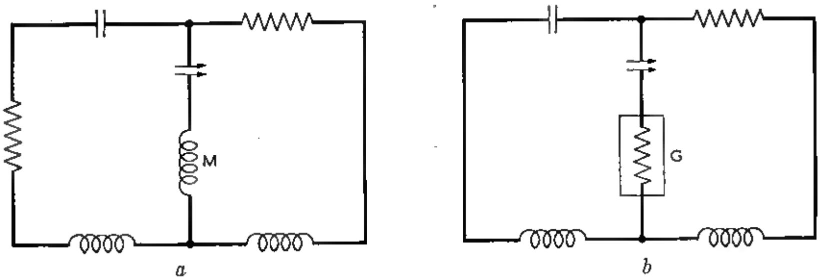


Fig. 5—Tuned Grid and Wind Reed Circuits.

In order to fix the rather useful concept of the analogy in mind, Fig. 5 is added showing the schematic circuit of the vacuum tube (5*a*) and a circuit diagram (5*b*), which represents determinant (31) the characteristic formulation of the dynamics of the larynx. Fig. 5*b* is represented by the conventions of an electrical circuit, except for the element G for which a different convention is necessary. The one taken here is that of a resistance enclosed in a rectangle. From (31) it will be seen to be similar to a resistance in its association with frequency p but different from resistance in that it occurs non-symmetrically in sign in the determinant. It does not involve dissipation.

Its occurrence here is the simplest possible for when there are appreciable concealed or ignored modes of motion it may have the form of a generalized impedance containing at least one element of resistance, but will always be non-symmetrical as a whole in sign in the determinant.

The use of the circuit for representing the mechanical system is an extension of an old but recently popularized method of studying mechanical or electrical vibrating systems by the help of analogy, one with the other. The extension consists in the explicit representation by diagram of the gyrostatic mutual which makes the determinant unsymmetrical in sign and of the unilateral mutual which makes the determinant unsymmetrical in magnitude. Fig. 6 is a diagrammatic representation of the more general system of determinant (30). This includes the external Y_0 reactions as well as the internal.

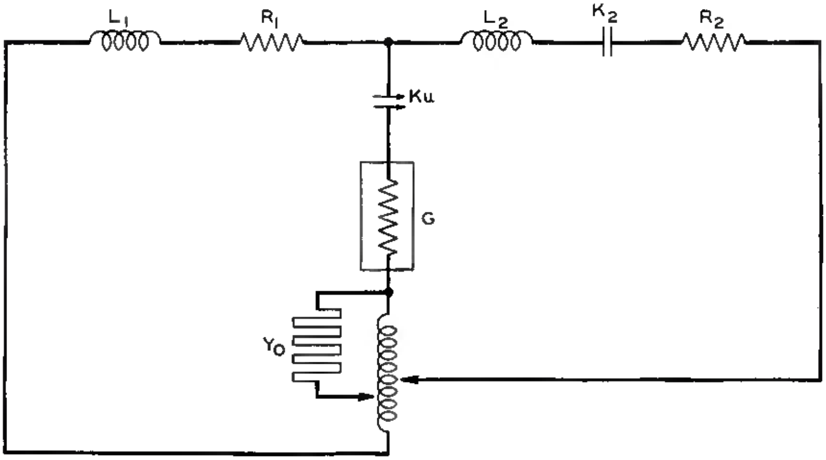


Fig. 6—General Wind Reed Circuit.

Having thus described the extended method of analogy the following study of the larynx with the help of the circuit diagram of its determinant should be clear.

SUSTAINED VIBRATION OF THE SIMPLE LARYNX

In vibrating, the vocal cords do not receive excitation of the frequency at which they vibrate. The source of power is in the air stream I_1 which enters the equations in K_u, the unilateral mutual impedance. Since this is treated as a constant circuit or dynamical element this air stream may be ignored as a drive and the resulting vibration considered as the free oscillation of the system. The determinant (31) (or 30) is then used to determine the free frequencies and decrements of the system. The method is as usual to solve for p in the equation

$$D = 0. \tag{33}$$

To simplify the demonstration the simple larynx without the load of the air chambers will be considered. Taking D of (31) then and expanding:

$$p^4 L_1 L_2 + p^3(L_1 R_2 + L_2 R_1) + p^2(L_1 K_2 + R_1 R_2 + G^2) + p(R_1 K_2 + G K_u) = 0. \tag{34}$$

If this be divided by $L_1 L_2$ and the uncoupled decrements and natural frequency defined:

$$\frac{R_1}{2L_1} = \Delta_1; \quad \frac{R_2}{2L_2} = \Delta_2; \quad \frac{K_2}{L_2} = \omega_2^2, \tag{35}$$

then

$$p\left[p^3 + p^2(2\Delta_1 + 2\Delta_2) + p\left(\omega_2^2 + 4\Delta_1\Delta_2 + \frac{G^2}{L_1 L_2}\right) + \left(2\Delta_1\omega_2^2 + \frac{G K_u}{L_1 L_2}\right)\right] = 0. \tag{36}$$

One of the roots of this equation is zero and another is negative real since all coefficients are positive. This root is therefore the decrement of a mode of non-vibratory motion. The remaining two roots may be real, imaginary or generally complex, of the form

$$\Delta \pm j\omega. \tag{37}$$

If it is found that $\Delta = 0$, then an oscillation once started will be sustained. If Δ be negative then any existing oscillation must subside or if Δ be found positive then an impulse will start an oscillation which of itself increases in amplitude to a point where its violence modifies the constants to such an extent as to make Δ vanish, leaving a sustained oscillation, or negative leaving the oscillation to subside to a lower amplitude or completely if sufficient permanent changes have been made.

If now (36) be written

$$Ap^3 + Bp^2 + Cp + D = 0 \tag{38}$$

and the first root (37) be substituted, two equations result, one from the real and the other from the imaginary terms, as follows:

$$A\Delta(\Delta^2 - 3\omega^2) + B(\Delta^2 - \omega^2) + C\Delta + D = 0, \tag{39}$$

$$A(3\Delta^2 - \omega^2) + B(2\Delta) + C = 0. \tag{40}$$

Now the condition for sustained oscillation is that $\Delta = 0$ and if the value of ω when this obtains be ω_0 then

$$\omega_0^2 = \frac{D}{B} \qquad \text{and} \qquad \omega_0^2 = \frac{C}{A}, \tag{41}$$

or the condition for sustained vibration in terms of the constants is

$$AD = BC. \tag{42}$$

In addition to this if use be made of the fact that in an algebraic equation such as (36) the coefficient of p^2 is the negative sum of all the roots then this coefficient is the real root. Let this be Δ_0 and then (36) may be written

$$p^3 + p^2\Delta_0 + p\omega_0^2 + \Delta_0\omega_0^2 = 0. \tag{43}$$

The coefficient of p in (36) is therefore the square of radial frequency at which sustained oscillation will take place and this is seen to be higher than the natural frequency ω_2 of the vocal cords, the difference being increased when the damping of either mesh is greater or when the coupling mutual G is greater.

It might be noted in passing that (43) is the free oscillation equation for any system which may be represented by a cubic equation and is not confined to the simple larynx. Such an equation always results when there is only one kind of reactive element in one of the meshes. It holds also for the tuned grid circuit.

The condition for sustained oscillation to be fulfilled for the constants may from (42) be reduced to:

$$R_1R_2 = G^2\left(\frac{L_1K_u}{GR_1} - 1\right). \tag{44}$$

It is rather difficult to place a simple physical interpretation on this formula. The qualitative import of it may however be seen by substituting the values of G and K_u from (13) and (10):

$$R_1R_2 = L_1^2I_1^2\left[\frac{\partial_0 R/\partial q_2}{R_1} - \frac{\partial_0 L_1/\partial q_2}{L_1}\right]\frac{\partial_0 L_1/\partial q_2}{L_1}. \tag{45}$$

The first term in brackets is in the nature of a resistance modulation constant, a fractional change in glottis resistance per unit displacement of the cords, to be designated by r and the second term similarly a glottis mass modulation constant, l. The quantity L_1I_1 is the momentum of the air in the glottis. This equation is then

$$R_1R_2 = (L_1I_1)^2(r - l)l. \tag{46}$$

Thus it appears that the resistance modulation must always be greater than the mass modulation and when the difference is small the air momentum must be increased to compensate. Owing to the physical limitation in accuracy of continuous maintenance of adjustment in the larynx, if a large momentum is depended upon to compensate for a small modulation difference, an unsteadiness or instability is likely to result. It is common experience that it is impossible to produce a sound with the voice with less than a certain minimum intensity. This corresponds, with the most favorable adjustment of the modulation constants which are physically possible, to a minimum momentum of air from the lungs which satisfies (46). It will be evident that this interpretation must not be taken too seriously quantitatively.

SUBSIDENCE AND ACCRETION OF VIBRATION OF THE SIMPLE LARYNX

Oscillograms made of the speaking voice show that, among other things, the amplitude of the oscillation and the pitch are in a continuous state of change. This is also true in singing but not nearly to the same extent. It seems therefore that in singing the adjustment of the voice system for sustained oscillation as described in (44) above is of major importance, while in speaking conditions for variation are of most importance.

The principle of the investigation of variation is simple enough but in all but the most elementary systems the algebra involved is impracticably awkward. If by solving (33) directly for the roots of p, it be found that Δ is positive, then any existing vibration will tend to increase while if Δ is negative, then vibration will tend to subside. The algebraic difficulties arise in the general solution but these are largely obviated by making the assumption, which is most likely usually fulfilled in practice, that the real parts of the roots may be treated as small quantities when compared with the imaginary parts. A common frequency for a man's voice is 150 cycles per second for which ω_0 is 1000 in round numbers. The decrement of a telephone receiver is ordinarily 100 to 200 in open air. The decrement of a tuning fork is represented by a fraction. Judging from variations in amplitude in an oscillogram (from which of course decrements may not be read directly) it would seem reasonable to assume that Δ is small compared with ω_0. The study of variation thus becomes an investigation of small departures from a condition of sustained oscillation, the reference condition being that critical adjustment for which the roots of interest of (33) are pure imaginary.

Suppose in (38) that $A = 1$; then without loss in generality:

$$p^3 + Bp^2 + Cp + D = 0. \tag{47}$$

In such an equation the roots are continuous functions of the coefficients. The same is true of their derivatives except at the one point where transition occurs from pure real to complex values. The values of the roots of interest in this connection are in their complex region at the point where the real part of the root passes through a zero value. This is the point at which free oscillation of the oscillating mode occurs, the values of the roots of this mode being as shown before, $\pm j\omega_0$.

If it now be supposed that one cause produces small variations, directly or indirectly on each of the coefficients and that the magnitude of this cause be x, then:

$$(3p^2 + 2Bp + C)\frac{dp}{dx} + p^2\frac{dB}{dx} + p\frac{dC}{dx} + \frac{dD}{dx} = 0. \tag{48}$$

The problem then is to determine dp resulting from any assigned cause dx when $p = j\omega_0$. From (43) we have at this point $B = \Delta_0$, $C = \omega_0^2$ and $D = \Delta_0\omega_0^2$.

$$2\omega_0^2\left(1 - j\frac{\Delta_0}{\omega_0}\right)\frac{dp}{dx} = -\,\omega_0^2\frac{dB}{dx} + j\omega_0\frac{dC}{dx} + \frac{dD}{dx}\,. \tag{49}$$

This is the frequency (complex) variation equation taken in the neighborhood of free oscillation.

When any readjustment of the larynx takes place all of the "constants" entering the coefficients undergo change, in particular those of the glottis K_u, R_1, G. Suppose for simplicity that one only varies, then this variation dK_u, dR_1, or dG may be taken as the magnitude of the cause dx. In particular if K_u vary,

$$dB = 0 = dC \qquad \text{and} \qquad dD/dx = G/L_1L_2,$$

$$\left(1 - j\frac{\Delta_0}{\omega_0}\right)\frac{dp}{dK_u} = \frac{G}{2\omega_0^2L_1L_2}\,. \tag{50}$$

If in addition Δ_0 be small compared with ω_0,

$$dp = \frac{GdK_u}{2\omega_0^2L_1L_2}\left(1 + j\frac{\Delta_0}{\omega_0}\right)\,. \tag{51}$$

This shows that if a condition of sustained oscillation is departed from by slightly increasing K_u, an increase in the amplitude of vibration begins which is proportional to the logarithm, since $(p + dp)$

is the exponent, of the increment dK_u and the frequency (imaginary part) of vibration increases slightly in proportion. If K_u were the only varying element the vibration would continue indefinitely to increase.

If on the other hand K_u be assumed constant, the variation being in R_1, then it may be similarly shown that

$$dp = \frac{d\Delta_1}{2\omega_0^2}\left[-\left(4\Delta_1 + \frac{G^2}{L_1L_2} + \Delta_0^2\right) + j\frac{\Delta_0\omega_2^2}{\omega_0^2}\right], \tag{52}$$

whence it appears that a small increase in glottis resistance dR_1 (or $d\Delta_1$) introduces a subsidence of vibration but an increase in frequency of oscillation as before. A decrease $-dR_1$ of course produces the opposite effect.

If the change be in G, it turns out that

$$dp = \frac{dG}{2\omega_0^2 L_1 L_2}\left[\left(K_u - 2G\Delta_0\right) + j\left(\frac{K_u\Delta_0}{\omega_0} + 2\omega_0 G\right)\right]\cdot \tag{53}$$

Here it appears that an increase in the gyrostatic mutual, G, may introduce either a subsidence or an accretion in amplitude but like the others makes for an increase in frequency of oscillation.

Variation in other elements produces similar conflicting tendencies not only in damping but in frequency.

The physical picture to be drawn from this is that in speaking the voice modulates from one amplitude and frequency to another by proper relative variations in adjustments in its constants, being constantly in a state of changing subsidence or accretion. It would seem also that the principal cause of change in frequency is in the vocal cords and that of amplitude variation in the glottis. Speaking is, in this respect, a more intricate process than singing.

Other Types of "Feed Back"

The detailed study of the larynx has so far been limited to the assumption that the "feed back" is entirely gyrostatic. This is of course actually not the case. How much influence is exerted by the general Y_0 is difficult to estimate.

If the trachea were a long tube but still shorter than a quarter wave-length of sound at the frequency of oscillation and rather smaller in diameter, and substantially open at the end the mass of the air in it would then be appreciable and Y_0 in (30) would be written p^2L_0. If in addition the gyrostatic term were negligible the system would then be exactly analogous with the tuned grid system and (32) rather than (31) should be the subject of detailed study.

If on the other hand the lungs acted substantially as a solid walled chamber of comparatively small size, the elasticity of the contained air would be represented by taking K_0 for Y_0. The surface area in the lungs is very large compared with a regular chamber of equal volume so considerable dissipation must be encountered by vibration. If this were the most important reaction Y_0 should have been replaced by pR_0.

Unquestionably all three types of reaction enter. A more general treatment to include them is plainly not a subject for a short paper. It is interesting however to note that in the dynamical system of brass horns these latter Y_0 reactions exert controlling influences. In this case the lips of the player perform the same function as do the vocal cords of the voice while the external load, the horn, corresponds to the pharynx, the reaction of which is the same dynamically as the trachea. In this case the frequency of the horn is that of sustained oscillation and not that of the lips. The same is true of the woodwind, in which case the reed or reeds replace the lips or vocal cords. In these cases Y_0 is proportional inversely to the hyperbolic tangent of the frequency or may be approximately represented by the impedance of an anti-resonant element.

Abstracts of Technical Articles From Bell System Sources

Notes on the Effect of Solar Disturbances on Transatlantic Radio Transmission.[1] CLIFFORD N. ANDERSON. In 1923 when the relation between abnormal long-wave radio transmission and solar disturbances was first noted, the outstanding abnormality was the great decrease in night time signal field strength accompanying storms in the earth's magnetic field. There was a slight increase in daylight signal field but this was distinctly secondary to the effect upon night field. Previous to 1927, data on signal fields were limited to one set of measurements a week, and although daylight signal field strengths were higher during periods of increased magnetic activity, it was somewhat difficult to determine the effect of individual storms. The present notes show the effects of individual storms of 60-kc transatlantic radio transmission and also give some indication as to their effect on short-wave radio transmission.

The Mutual Impedance Between Adjacent Antennas.[2] CARL R. ENGLUND and ARTHUR B. CRAWFORD. The simple theory for the computation of reflecting or multibranch antenna systems is sketched. If the points at which observations of electrical quantities are to be made are definitely specified, a knowledge of the self and mutual impedances (properly defined) between antennas is sufficient to make the computations determinate. Of the circuit constants, the most useful and accessible is the antenna current ratio

$$K_{12} = \frac{I_2}{I_1} = K_0 e^{i(\phi - (2\pi d/\lambda))}$$

and in the work here reported ϕ has been measured in the range 0.33 λ to 1 λ. Experiment has shown that in this range ϕ is that theoretically calculable for a Hertzian doublet. Actually this range is equivalent to $\lambda/3$ to ∞. The discussion of experimental procedure is purposely thorough.

An Experimental Method for the Determination of the Ballistic Demagnetization Factor.[3] DONALD FOSTER. A method is described for experimentally determining the ballistic demagnetization factor. By means of a double search coil of novel design the magnetization and

---

[1] *Proceedings of the Institute of Radio Engineers*, September, 1929.

[2] *Proceedings of the Institute of Radio Engineers*, August, 1929.

[3] *Philosophical Magazine*, September, 1929.

the magnetic field intensity are determined from ballistic galvanometer deflections. While the discussion refers mainly to circular cylinders, the scheme is adaptable to specimens of other shapes. It is particularly designed to obtain accurate measurements of field intensity in cylinders of small diameter.

Details of a special design are given.

Curves are given which illustrate the variation of the demagnetization factor with the magnetization, as well as the dependence of this relation on the material and on the dimensional ratio.

The Use of Continued Fractions in the Design of Electrical Networks.[4] THORNTON C. FRY. In U. S. Patent No. 1,570,215 and in several technical papers by Bartlett and Cauer it has been shown that continued fractions can often be used in designing networks with preassigned impedances. The chief difficulty of the method has been that it frequently required the structures to contain negative resistances, inductances or capacities and therefore the results, though correct in theory, were often worthless in practice because the networks could not be constructed.

The present paper removes this difficulty in virtually all cases where the analytic character of the desired impedance is known, that is, where it can be represented by a formula and not merely by a graph. In such cases the choice of a type of structure, as well as the assignment of values to the elements, becomes almost a matter of routine with the definite assurance in advance that no negative elements will be required.

A Voltage Regulator for Gas Discharge X-Ray Tubes.[5] F. E. HAWORTH. This note describes a device used in connection with a gas discharge x-ray tube, to regulate the voltage across it by automatically adjusting a mercury valve between the tube and the pumps, thus controlling the pressure of the gas. It has been used with tubes of the Hadding and Shearer types and has operated satisfactorily for more than a year. It was designed to replace the regulator described by Bozorth, which is similar in principle but has certain disadvantages, for example the moving parts have high inertia and adjustment is required when the atmospheric pressure changes.

The Significance of the Hydrogen Content of Charcoals.[6] H. H. LOWRY. Most studies of the thermal decomposition of hydrocarbons

---

[4] *Am. Math. Soc. Bull.*, July–August, 1929.

[5] *Journal of the Optical Society of America*, August, 1929.

[6] *Journal of Physical Chemistry*, September, 1929.

are confined to an examination of the composition of the liquid and gaseous products. Among exceptions to this generalization may be mentioned the interest in coke, carbon black, and charcoal. Even in these cases the physical properties rather than the chemical composition are regarded as the factors which determine their suitability for specific uses. However, in an earlier paper it was pointed out that certain physical properties of a group of charcoals were rather simply related to the per cent hydrogen which was contained in them as determined by ultimate analysis. This group of charcoals was prepared in a gas-fired furnace from a single, specially-selected lot of anthracite coal. As stated in this earlier paper, careful consideration of the commercial records taken at the time of preparation indicated that the hydrogen content was probably determined by the maximum temperature to which the samples were heated during their preparation. The hydrogen contents ranged from 0.21 to 0.53%, while the probable range of maximum temperature was 900° to 1200°. The presence of hydrogen in these charcoals was shown to be consistent with a point of view that so-called "amorphous" carbons are hydrocarbons of low hydrogen content built up of polymerized residues from the thermal decomposition of hydrocarbons of greater hydrogen content. Since the significance of the hydrogen content of charcoals has been generally overlooked, the present study was undertaken in order to evaluate the factors which may ordinarily be varied in the preparation of charcoals for various purposes. The factors which were independently varied in this study were the maximum temperature, the time of heating, the atmosphere surrounding the sample during heating and the raw material. To a limited extent the effect of previous heat treatment was also determined. A later paper will give the results of the study of the correlation of hydrogen content and some adsorptive properties of charcoals prepared under carefully controlled conditions.

Beginnings of Telephony.[7] FREDERICK LELAND RHODES, Outside Plant Development Engineer, Department of Development and Research, American Telephone and Telegraph Company.

It is only within the past decade or so that science and business have become subjects for literature. Somehow these great phases of human endeavor have been sadly neglected in the literary world until very recently, and now it seems as though, conscious of the lack of good literature in these fields, engineers, scientists and business executives are making up for lost time. Frederick Leland Rhodes has written a new book which undoubtedly will be of great assistance to those in the

[7] Harper & Brothers, New York and London, 1929.

telephone industry, for it supplies them with an accurate picture of the technical background of a great industry. It is greatly to the advantage of an individual to know the history of his own business, and Mr. Rhodes has supplied it in an interesting form, thoroughly accurate and readable. No effort has been made to set down the more recent achievements in the world of telephony, but only to carry each chapter to what might be termed the "middle period" in development. There are many phases of the telephonic art which have not been touched upon in the volume, but at the same time, one is not conscious of any lack in this respect as one reads through its interesting pages.

Any volume is the better off for illustrations, and Mr. Rhodes' book is generous in that it carries fifty-four illustrations scattered through 260 pages.

The first portion of the book naturally deals with Alexander Graham Bell and occupies three chapters. Following this we have two chapters called "The Bell Patents." As General John J. Carty, Vice President of the American Telephone and Telegraph Company, says: "Never before had the claims of an inventor been subjected to such exhaustive litigation and judicious scrutiny, and never before did an inventor receive such a complete and dramatic vindication." The remainder of the fourteen chapters deals with the truly romantic progress of telephone plant, its improvements and expansion over a term of years when telephony was young and the road was fraught with difficulty. Of special interest are the numerous references to original and authentic sources, and in this regard the author has unquestionably used great care and much labor in order to give his reader the most accurate information possible, thus more truly gaining his end of supplying a concrete picture of the younger days of a great industry.

Mr. Rhodes' volume is a great contribution, not only to the literature of telephony, but also to that rapidly growing library which contains in its pages the romance of business in America. As a library reference book it will be valuable to the technical student. Any member of the Bell System would do well to familiarize himself with this work, not only because it will help him in his job, but because he will find it a really interesting story.

Further Note on the Ionization in the Upper Atmosphere.[8] J. C. SCHELLENG. In this paper Mr. Schelleng records certain considerations that were omitted from a previous paper, which omission resulted in some difficulty.

---

[8] *Proceedings of the Institute of Radio Engineers*, August, 1929.

Transmission Networks and Wave Filters.[9] T. E. SHEA. In this book is summarized the research and experience of the Bell System in the application of electric wave filters, equalizers, balancing networks and similar electrical systems. The preface discusses the nature of the signals transmitted over communication systems and a statement of the principal ways in which selective networks are used to modify signal transmission. A detailed example of the application of selective networks to an actual long distance telephone circuit gives specific engineering requirements and limitations.

The next portion of the book deals with some of the more general principles governing network analysis. The engineering terms used to evaluate network performance are described and a number of general theorems and equivalences which simplify the analytic treatment of networks are demonstrated. A considerable discussion is also given of the characteristics of the elementary two-terminal networks most used as constituents of larger structures.

With this background the author is now ready to consider the properties of wave filters. Conditions for free transmission and attenuation in ladder networks are set up and the particular networks of chief practical importance are described in detail. The various structures revealed by this listing differ widely among themselves as regards propagative and impedance characteristics even when they transmit the same frequency bands. Since the ideal network characteristics seldom correspond exactly to any one of these structures, filter requirements are usually met most efficiently by composite networks, containing sections of several different types. The author describes the conditions which must be satisfied before different sections are joined together and gives several examples of methods of computing the performance of such composite structures.

This treatment of networks deals only with their response to steady single-frequency electrical impulses. It cannot be applied directly to communication systems, since signals are of more complicated wave forms and are transient in character. In the last portion of the book therefore, the author discusses the use of Fourier analysis in relating the characteristic of the network computed on a steady-state basis to its response to a transient impulse of arbitrary character.

Some Principles of Broadcast Frequency Allocation.[10] L. E. WHITTEMORE. This paper discusses some of the technical factors which must be considered in the allocation of frequencies to broadcasting stations

---

[9] D. Van Nostrand Company, New York.
[10] *Proceedings, Institute of Radio Engineers*, August, 1929.

in such a way as to provide the best possible coverage of a given country or continental area.

A given frequency or channel can be used for either of two kinds of service; (1) by one station, exclusively, to give high grade service to the immediate locality and opportunity for service over broad rural areas when transmission conditions are good, and (2) by two or more stations simultaneously, to give local service to a number of separate regions, each of rather restricted area. The problem, therefore, involves a determination of (1) the proper balance between the two kinds of service, rural and urban, and (2) the proper basis for the apportionment of the assignments.

Reference is made to the basis of apportionment of radio broadcasting assignments laid down in the U. S. Radio Act of 1927, and to certain suggestions which have been made for the apportionment of broadcasting frequency assignments among the countries of Europe.

A brief discussion is given of the relation between field intensity, or signal strength, and distance of transmission at broadcast frequencies. The paper also discusses briefly the effects produced in the case of (1) a single station operating exclusively on a "clear" channel, and (2) two or more stations operating simultaneously on the same channel.

It is suggested that the distribution of assignments on "clear" channels, in a given continental area be made proportional to the population of each of several large geographical units or zones and that the distribution of assignments on "multiple assignment" channels be made to comparatively small geographical units in proportion to their areas.

Contributors to this Issue

John R. Carson, B.S., Princeton, 1907; E.E., 1909; M.S., 1912; American Telephone and Telegraph Company, 1914–. Mr. Carson is well known through his theoretical transmission studies and has published extensively on electric circuit theory and electric wave propagation.

A. B. Clark, B.E.E., University of Michigan, 1911; American Telephone and Telegraph Company, 1911–. Toll Transmission Development Engineer, 1928–. Mr. Clark's work has been largely concerned with toll telephone and telegraph systems.

Karl K. Darrow, B.S., University of Chicago, 1911; University of Paris, 1911–12; University of Berlin, 1912; Ph.D., University of Chicago, 1917; Western Electric Company, 1917–25; Bell Telephone Laboratories, 1925–. Dr. Darrow has been engaged largely in writing on various fields of physics and the allied sciences. Some of his earlier articles on Contemporary Physics form the nucleus of a recently published book entitled "Introduction to Contemporary Physics" (D. Van Nostrand Company). A recent article has been translated and published in Germany under the title "Einleitung in die Wellenmechanik."

Bancroft Gherardi, B.Sc., Polytechnic Institute, Brooklyn, N. Y., 1891; M.E., Cornell University, 1893; M.M.E., Cornell University, 1894. New York Telephone Company, Engineering Assistant, 1895–99; Traffic Engineer, 1899–1900. New York and New Jersey Telephone Company. Chief Engineer, 1900–06. New York Telephone Company, and New York and New Jersey Telephone Company, Assistant Chief Engineer, 1906–07. American Telephone and Telegraph Company, Equipment Engineer, 1907–09; Engineer of Plant, 1909–18; Acting Chief Engineer, 1918–19; Chief Engineer, 1919–20; Vice President and Chief Engineer, 1920–. Mr. Gherardi is a Past President of the American Institute of Electrical Engineers.

Frank B. Jewett, A.B., California Institute of Technology, 1898; Ph.D., University of Chicago, 1902. American Telephone and Telegraph Company, Transmission and Protection Engineer, 1904–12. Western Electric Company, Assistant Chief Engineer, 1912–16; Chief Engineer, 1916–21; Vice President and Chief Engineer, 1921–22; Vice President, 1922–25. International Western Electric Company, Vice President, 1922–25. Manufacturers Junction Railway, Vice President, 1922–25. American Telephone and Telegraph Company, Vice President, and Bell Telephone Laboratories, President, 1925–. Dr. Jewett is a Past President of the American Institute of Electrical Engineers.

Francis F. Lucas, Associated Bell Telephone Companies, 1902–10; Western Electric Company, 1910–25; Bell Telephone Laboratories, 1925–. Mr. Lucas has specialized in the development and application of microscopy. He has received international recognition and awards for the development of high power metallography and ultra-violet microscopy and for numerous scientific papers which he has contributed on the subjects of metallurgical and biological research. For several years he has been Consulting Technical Expert for the War Department, U. S. A., Watertown Arsenal.

Edward L. Nelson, B.S. in E.E., Armour Institute of Technology, 1914; Western Electric Company, 1917–25; Bell Telephone Laboratories, 1925–. As Radio Development Engineer of Bell Telephone Laboratories, Mr. Nelson is responsible for the development and design of commercial radio apparatus, which includes radio broadcasting equipment.

R. L. Wegel, A.B., Ripon College, 1910; Assistant in Physics, University of Wisconsin, M.A., 1910–12; Western Electric Company, 1914–25; Bell Telephone Laboratories, 1925–. Mr. Wegel has written several papers on theory of telephone receivers and on the theory of hearing. The article appearing in this issue is taken from lecture notes on Mechanics of Vibrating Systems by the author. It is planned to publish these notes in future issues of the Bell System Technical Journal.

M. K. Zinn, B.S. in E.E., Purdue University, 1918; American Telephone and Telegraph Company, 1919–. Mr. Zinn's work has been related particularly to the design of loading for telephone circuits.

CORRECTION SLIP FOR ISSUE OF JANUARY, 1930

Page 153: Equation (10) should read

$$h(t) \sim \left\{ 1 + \frac{1}{1!}\left(\frac{1}{4\lambda t}\right) + \frac{1^2 \cdot 3^2}{2!}\left(\frac{1}{4\lambda t}\right)^2 + \cdots \right\} \frac{1}{\sqrt{\pi \lambda t}},$$

$$h(t) \sim S(\lambda t) \tag{10}$$