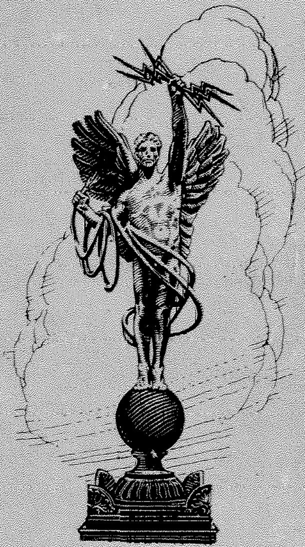
Mr Pocock

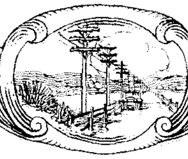
Eng. Laborator

ELECTRICAL COMMUNICATION



NOVEMBER
1922

No. 2
VOL. 1

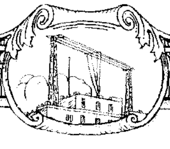


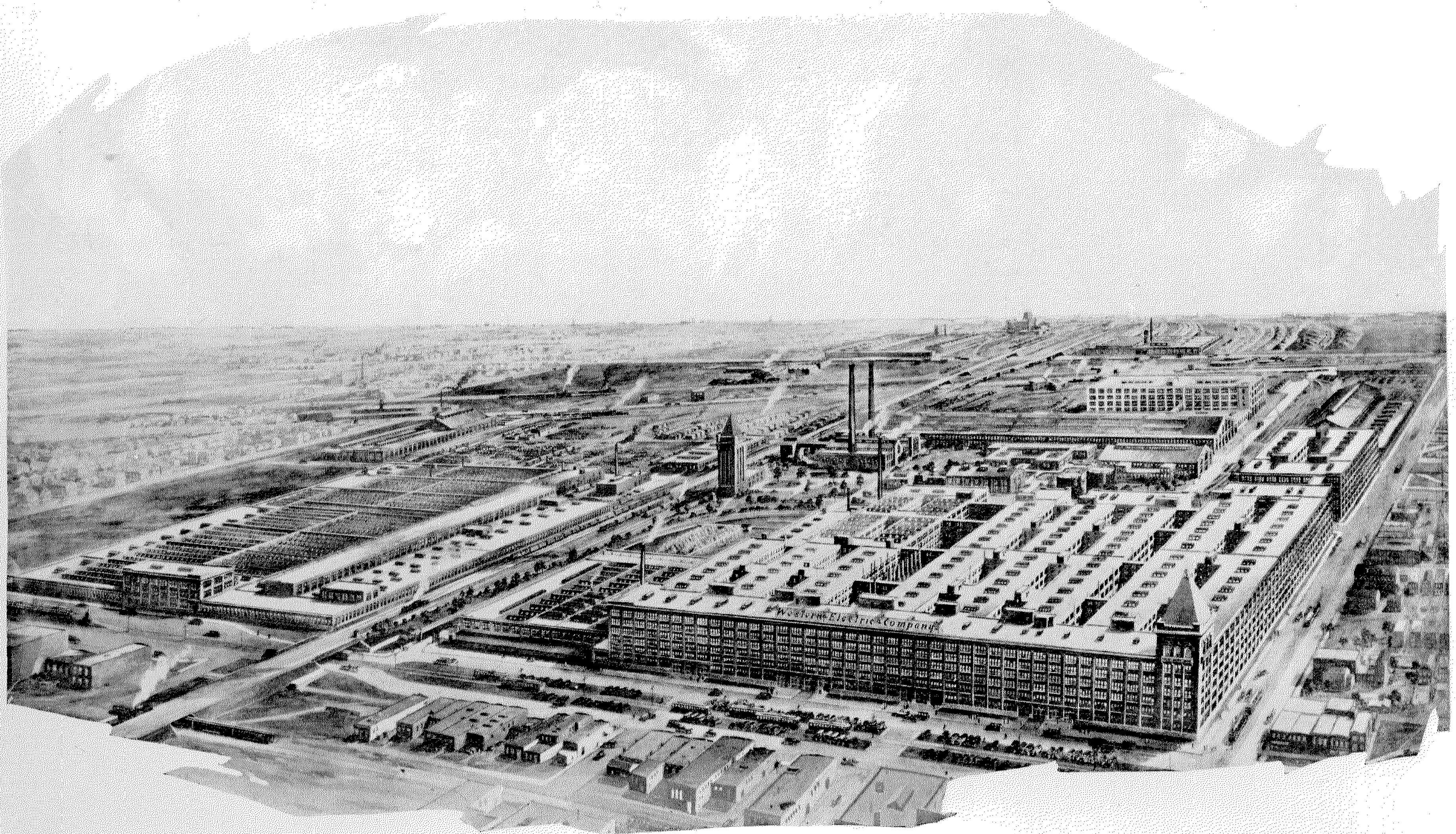
ELECTRICAL COMMUNICATION

A Journal of Progress in the
Telephone, Telegraph and Radio Art
Published Quarterly by the
INTERNATIONAL WESTERN ELECTRIC COMPANY
INCORPORATED
195 BROADWAY, NEW YORK, N. Y., U. S. A.
Charles G. DuBois, President George C. Pratt, Secretary J. W. Johnston, Treasurer
Subscription $1.50 per year; single copies 50 cents

Volume I NOVEMBER, 1922 Number 2

CONTENTS





Hawthorne, Chicago, Works of the Western Electric Company

The Hawthorne Plant of the Western Electric Company

By H. F. ALBRIGHT

Vice-President and General Superintendent, Western Electric Company

IT is probably a familiar fact to many that the Hawthorne Works of the Western Electric Company, is admittedly the world's largest manufactory of telephones and telephone apparatus. To give a concrete idea of just what its bigness means is somewhat difficult. A landowner who can visualize definitely the extent of an acre, will realize the significance of the statement that the floor space in the buildings aggregates eighty acres, but even such an expression does not convey an adequate impression of the imposing spread of the Hawthorne buildings, forming as they do a notable landmark on the western outskirts of Chicago.

A simple statement that the Works uses over 1,500 million feet of wire in making lead covered telephone cable in the course of a month conveys to most persons merely an indefinite sense of vastness. Such a statement may be made more definite if expressed in larger units and shorter periods, as, for example, that the cable plant uses approximately 12,000 miles of wire for an average day's output of lead covered cable. This wire is insulated with continuous spirals of paper ribbon amounting to 350 tons per month, while the lead used for the sheath approximates 150 tons for an average day's output.

Notwithstanding the magnitude of cable manufacture, as indicated by these figures, it by no means forms the major part of Hawthorne's activities. The manufacture of telephones and switchboard equipment, train dispatching telephone equipment, printing telegraph, and other apparatus, together covering practically the entire field of electrical communication, requires a much large proportion of space than is utilized in the manufacture of cable. In the switchboard shop there is built every type of telephone switchboard from the small magneto to the largest central battery central office types including the more recently developed machine switching central office types. There are produced in an average year at Hawthorne more than 9,000 sections of telephone switchboard, 1,800,000 telephone receivers and transmitters, 4,000,000 relays, 10,000,000 jacks, and 6,500,000 cords.

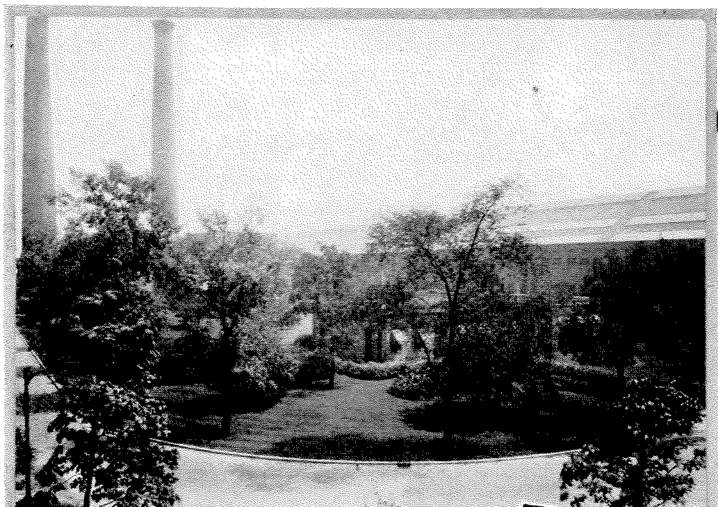
Over 10,000 different codified types of telephone, telegraph and radio apparatus are manufactured. These 10,000 different types of apparatus require in their completed form more than 100,000 different pieces or component parts. There is an average of twelve machine or manual operations performed on each part, resulting in a total of 1,200,000 operations. Some of these types of apparatus are produced in quantities of half a million per month. These statements may help to give some conception of the enormous amount of work passing through the factory and to explain why there are required buildings having a floor area of an eighth of a square mile.

The first group of Hawthorne factory buildings was erected in 1904–1905 on what was then a barren tract of 262 acres of prairie land. They provided approximately 18½ acres of floor space. Since this first group of buildings was erected the factory has grown at an average rate of over 3½ acres of floor space per year. The growth still continues and at present there is under construction a rolling mill for producing wire rods, a wire mill and an enlarged woodworking department.

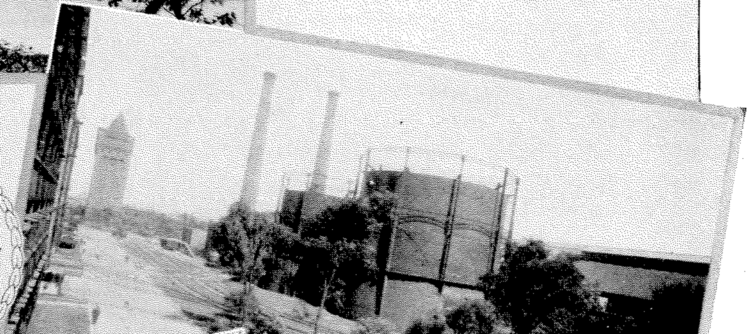
Practically all of the manufacturing buildings are built on a uniform architectural plan, so that departments can be very easily regrouped when required by the growth of the business or other considerations. In a few instances, as in the case of cable manufacture, a more specialized form of building is constructed, due to the peculiar demands of the class of work involved.

The main manufacturing group consists of long parallel buildings arranged at right angles to main buildings which connect them together at the ends. The buildings are sixty feet in width which insures an ample supply of natural light from the numerous windows that occupy practically all of the side walls. Connecting wings tie the parallel buildings together across the intervening light wells, and furnish space

Around the Hawthorne Works



Works Hospital
with a glimpse of
surrounding landscape
gardening effects



Gas
Producer
Plant

Power
Plant

✠

Cable Plant
A view down the
big lead press room

✠

for elevators and stairways, thus leaving the manufacturing areas unobstructed. Fire walls separate the buildings into sections each containing approximately 15,000 square feet of floor area. Each section is provided with at least three exits equipped with double fire doors which close automatically in case of fire.

All of the buildings are of fireproof construction and equipped with a very extensive automatic sprinkler system. No space in any building is more than fifteen feet from an automatic sprinkler head, forty feet from a water pail, sixty-five feet from a rack of 1½ inch fire hose and seventy-five feet from a rack of 2½ inch fire hose. Whenever an automatic sprinkler head is actuated, it signals fire headquarters and indicates its location on an annunciator. A trained fire department which is on duty night and day responds immediately. Numerous manual fire alarm boxes located throughout the Plant also connect with fire headquarters. Constant cleanliness and the prompt removal of all rubbish minimize fire hazards. Wide unencumbered aisles insure easy egress and frequent fire drills train the workers to evacuate any section promptly and without confusion.

A great deal of attention and study have been expended in making the factory buildings comfortable and pleasant places in which to work. Very large window areas insure ample daylight normally, and when artificial light is necessary it is supplied from numerous lighting fixtures, efficiently located. The floors are swept each day and in addition are scrubbed frequently, and in other respects the shops as well as the offices are kept orderly and clean. Drinking water which is filtered and cooled is supplied from many sanitary drinking fountains.

In addition to the efforts which have been put forth to make the inside of the buildings sanitary, comfortable and attractive, much time and thought have been devoted to the surroundings. The Hawthorne Works stand on what was originally a bare tract of yellow clay. Since the beginning in 1905, this plot has been transformed by trees, shrubs, flowers, broad cement roadways and sidewalks. The surface clay had to be removed and replaced by fertile loam to make this landscaping possible, but the results have well repaid the effort. It is sometimes hard for visitors to believe that the place has not been under cultivation for many years when they see the numerous large trees which have been moved to their present locations. These trees at the time of transplanting in many instances weighed from ten to twelve tons.

Nothing unsightly is permitted to disfigure the grounds. All pipes, wires, etc., are carried underground in large tunnels, 7½ feet high and 7 feet wide. Two miles of such tunnels connect the different buildings comprising the Works.

Steam and electric power used at the Works is generated in its own power house. Two and one-half million kilowatt-hours of electrical energy are used each month for power and light. The boilers generate 100,000,000 pounds of steam per month, and 100,000 tons of coal are burned during an average year.

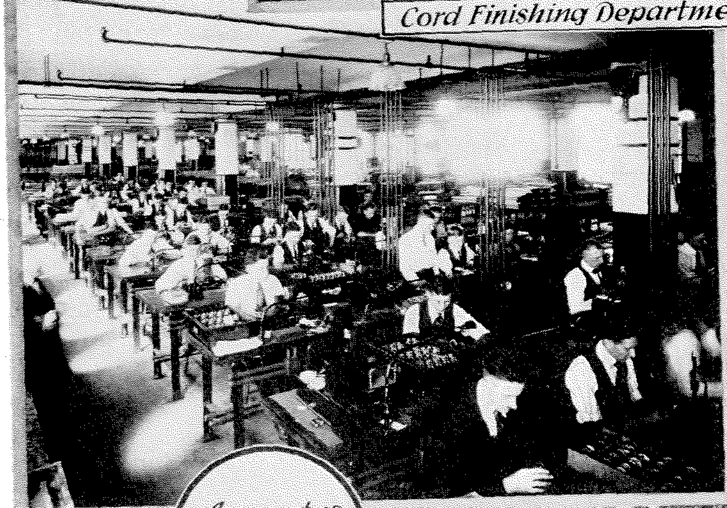
Hawthorne uses as much gas as a city of 100,000 inhabitants. This is not used for either heating or lighting purposes but entirely for industrial processes requiring high temperatures, such as annealing, hardening, core baking, glass making, metal bending, incandescent lamp manufacture, etc. Twenty-five million cubic feet of gas are used monthly. It is made in the Works Gas Plant by the producer process.

This excellent factory, with its thoroughly modern equipment, would be of little use, however, without its large body of skilled, intelligent artisans, competent executives and trained manufacturing engineers. Altogether, about 27,000 people now constitute the factory's population. Every one of these, directly or indirectly, plays some part in keeping Hawthorne's product up to Western Electric standards of excellence—the most exacting in the world.

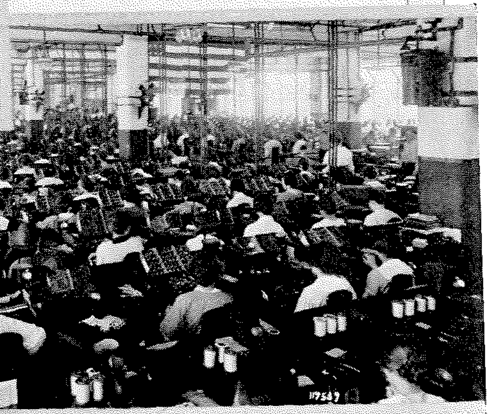
Telephone and telegraph apparatus is of more or less delicate construction and adjustment, and must therefore be manufactured with great care. The Western Electric requirement of perfect interchangeability of parts makes extremely close workmanship a factory necessity. It is a very simple matter to achieve quantity production of parts that are "not quite" interchangeable, or "almost" interchangeable, but if parts are to fit exactly in every instance, without requiring any filing, drilling, hammering or other work every detail of the apparatus must be held to extremely close limits. Real interchangeability means

Intimate views of Hawthorne Works

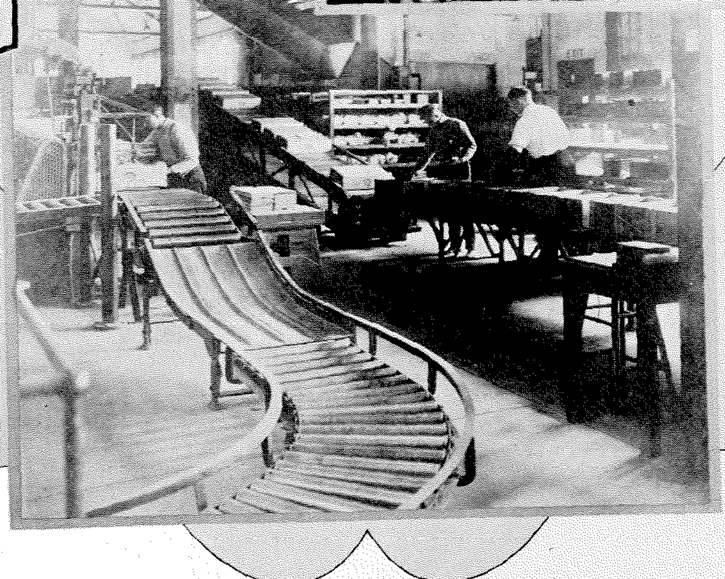
Cord Finishing Department



Apparatus Assembling Department



Coil Winding Department



Packing Department

Showing use of mechanical conveyors

that everything entering the manufacture of the apparatus must be right—the design, the raw material, the methods of manufacture, the workmanship, the tools and the setting of the machines. Many of the punch press parts produced, for example, are held to such close limits that a slight change in the composition of the brass they are made from will make it impossible to produce them accurately enough, even with perfect tools, perfect tool-setting and perfect manipulation.

Consequently the first care to insure Western Electric quality begins with the raw materials, all of which are carefully inspected to insure that they meet all the requirements of the raw material specifications. Some idea of the magnitude of the raw material inspection may be had when it is realized that there are used annually at Hawthorne, 20,000 tons of copper, 10,000 tons of steel, 45,000 tons of lead, 10,000,000 pounds of brass, 2,000,000 pounds of silk and cotton and 25,000,000 feet of lumber. Smaller quantities of many kinds of material are also required, over 20,000 varieties being carried in the raw material store rooms.

Rigid inspection through the manufacturing operations follows this initial raw material inspection. The inspectors check the tools, tool settings and workmanship at every stage, to insure the suitability and standardization of the parts entering into the assembled piece of apparatus.

Some of the tools used are marvels of accuracy. To secure extremely close limits on parts, the tools that make these parts must be held even closer to absolute dimensions than the parts themselves, since stretch of the metal used in their manufacture and other causes introduce unavoidable variations in size. The perforating punch and die that produces the commutator strips used in machine switching apparatus furnishes a good example of the fine tool-making necessary in modern telephone manufacturing work. Fifty punches, fifty dies and fifty shedders, mounted in suitable steel yoke plates, form the punch and die. Each of these fifty similar parts differs from the others by less than one one-hundred-thousandth of an inch. In fact, the variations in the sections are so small that if two of these punches and dies were completely disassembled, the parts could be put together in any combination and produce two tools which would not vary more than plus or minus two one-hundred-thousandth of an inch over the entire length of fifty sections. Bearing in mind that one one-hundred-thousandth of an inch is about one one-hundredth of the thickness of a piece of thin tissue paper, the extreme accuracy of such tools can be appreciated.

In addition to the process inspection of the component parts of the apparatus as they go through the machine operations, the finished apparatus is subjected to a rigid inspection to insure that it properly meets the performance specification in every particular, as well as fulfilling the Western Electric standards for quality and workmanlike appearance. As a final check, a fixed proportion of each class of apparatus is reinspected by an entirely different group of the Inspection Department, known as the Ultimate Inspection Department.

The result of all this extreme care in methods, tools, workmanship and inspection is the maintenance of Western Electric quality, noteworthy in the electrical industry for over fifty years.

The Future of Long Distance Telephony in Europe

By FRANK GILL, O. B. E.

European Chief Engineer, International Western Electric Company

(*This paper formed a part of the inaugural address by the President of the British Institution of Electrical Engineers at the Annual Meeting, held in London, November 2, 1922.* —EDITOR.)

I HAVE thought that the subject to which I could most usefully direct your attention is one relating to the art of Electrical Communication, and particularly, though not exclusively, to telephony over considerable distances. I propose, therefore, briefly to review some of the recent advances in the telephone art which affect long distance communication in Europe, or as I prefer to call it, "through communication," because distance is not necessarily a feature, though often it is present.

If we consider primitive man, his first and immediate need is for food, then shelter and defense, then tools and clothes; but directly he has arrived at the state in which his own and his immediate neighbors' wants have been supplied, so far as their own exertions can supply them, the need for communication arises, and that even before the need for transportation. There is, however, no need to insist on any priority as between these two arts; they are so intimately connected that we may, without any violence to meaning, define communication as transportation of intelligence; without communication man cannot know where to obtain such of his requirements as he himself is unable to satisfy; without it there is no use in his producing more than he requires for himself, since in its absence, he cannot know whence arises a demand for his spare produce. While all this is true of primitive man it has applied much more intensely since machinery came to the aid of production in the complicated system of trade which now serves the world. That which fifty years ago was regarded as a luxury to be enjoyed only by the few, is now a necessity to the many. Today no nation stands alone or is sufficient for itself; more and more the interdependence of nations is being recognized; and more and more is it realized that no nation can be prosperous or afflicted without a result being felt by other nations. It is easy to illustrate this community of nations by taking the clothes we wear, the food we eat, but all this is well known to you; it is sufficient to say that to bring together the products of all the world to your doors is the work of communication and tansportation, and that the efficiency of both is vital. Other things being equal, the nation best equipped with the means of production, communication and transportation will enjoy a great advantage in the race for commercial supremacy, and perhaps also in the search after national well-being. It follows therefore that a great responsibility is laid on those to whom is entrusted the means of communication, or who control those means, whether Government department, public company, or other agency, and particularly so because it is at last generally recognized that competition is not an aid to efficiency in this business. In return therefore for the grant of facilities to carry on its work, each grantee authority must ultimately recognize its duty to the public, and if it realizes this, dare not adopt either the selfish attitude of attempting to make as much profit as possible, nor the passive attitude of merely supplying the service demanded by the public. To discharge its duty it must diligently and actively search out new means and facilities and also set about educating the public in regard to their need for communication. The authority is the custodian of the knowledge, it must teach the public what the public did not at first realize—that efficient communication is the life blood of commerce and of national and international understanding and amity—and without effective communication there would appear to be little chance of success for such projects as the League of Nations. It is by this campaign of education that the well developed industries of the world have been stimulated, not by merely diligently serving the public demand; the railways have themselves created demands for traffic, the press has done the same, so has every great and successful industry, so must the telephone industry if it is not to fail in the performance of its duty.

The fact that so much of the means of electrical communication in Europe is under Goverment ownership, and under the present prevailing custom removed from the stimulus of profit earning, makes it quite possible that Governments may instruct their Telephone and Telegraph Departments that they do their duty when they diligently supply the public demands for service, and perhaps such a course is the more probable when the many other calls for expenditure experienced by Governments today are considered. But here a deliberate choice must be made; is communication and even better communication a necessity or not? If it is, then the passive attitude of merely satisfying public demands must be abandoned and an aggressive attitude take its place. It is not by a passive attitude that the great development in telephones has been built up in the United States, Canada, Denmark, Sweden and Norway, but rather by a resolute, purposeful and well directed campaign of education of the public and of existing users. In the United States particularly, there has for years been an educational campaign of a very high order, coupled with the construction of plant in advance, without which it is, of course, worse than useless to create demand. All this seems self-evident and trite, but two questions will test very quickly whether in fact this matter is quite so self-evident and obvious as it appears; these questions are—(1) Has telephony, during the 46 years it has been available, been of as much use to Europe as it might have been? (2) Have the organizations, Government and otherwise, been permitted to do what they have wished to do? The answer to both questions is most decidedly—No.

Attenuation-Frequency Characteristic of 1 Mile of 0.165 Inch Open Wire Line—
Dry Weather Conditions Assumed

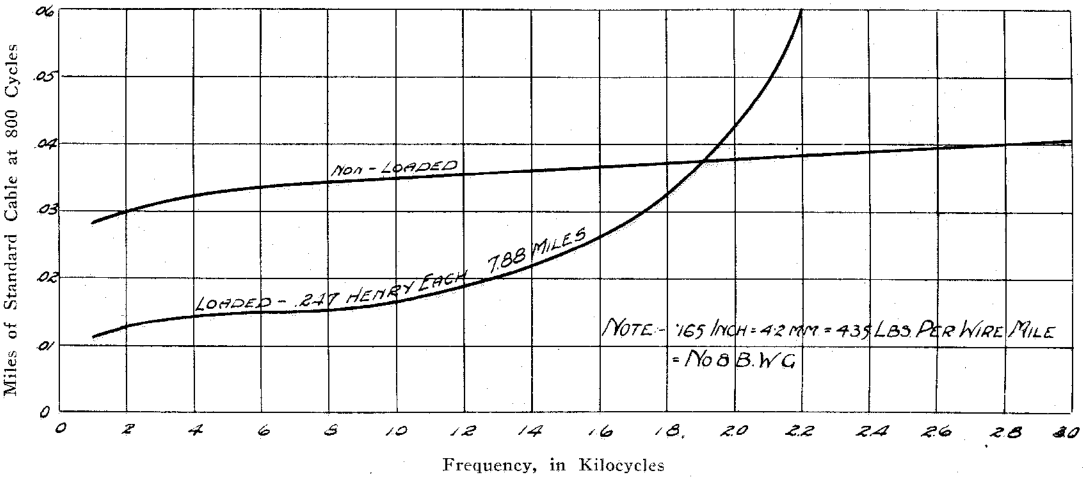


FIGURE 1

It will be noticed that above it was assumed, for the moment, that a Government department should not earn any more money than necessary to be self-supporting, but it seems also wrong to reason that a Government department should not earn something more than just enough merely to pay its way; it is undoubtedly healthy for the *esprit de corps* of any organization, that its personnel should know that their organization is an efficient one, in which they can take a justifiable pride, and one of the ways of testing the efficiency of public concerns is to associate service with returns. With a staff comprising many persons, it is unhealthy that the idea should prevail that profit earning is of no account. There is, however, more than the question of the effect on the staff, important though that be. Without a surplus of income over expenses, there is no margin for unforeseen contingencies which must constantly arise in such a flexible business; service trials and research are likely to be adversely acted upon and capital will be raised with greater difficulty. Further, there seems no reason why a Government should not include in the rentals a sum plainly intended to be a contribution towards revenue; it is difficult to see any reason why it is permissible, for the purpose of raising revenue, to tax, say, food, but not telephones, or why it is proper to make a considerable surplus on postage, but not on telephones. It would seem that the correct course is for a Government, if it operates the telephones of a nation, to raise from them something towards the National Revenue and pay such a return on the capital invested in the business as to make certain its ability to raise whatever money may be required to extend the business.

Let us now pass on to consider some of the alterations in practice caused by recent developments in telephony as they affect long distance or through communication.

LOADING

By this term is meant the deliberate addition of inductance to the circuit for the purpose of increasing the distance over which satisfactory speech is feasible. Such inductance may be in the form of evenly distributed inductance effected by wrapping the copper conductor with magnetic material, such as fine iron wire, or, and more commonly, it may be in the form of lumped inductances obtained by inserting in series in the circuit at intervals, coils having the required inductance and a minimum resistance. It was in 1887 that Oliver Heaviside pointed out that the addition of inductance by either of these methods, (inductance being then thought harmful) would be beneficial to the transmission of speech.

Attenuation-Frequency Characteristic of 1 Mile of .036 Inch Cable Circuit

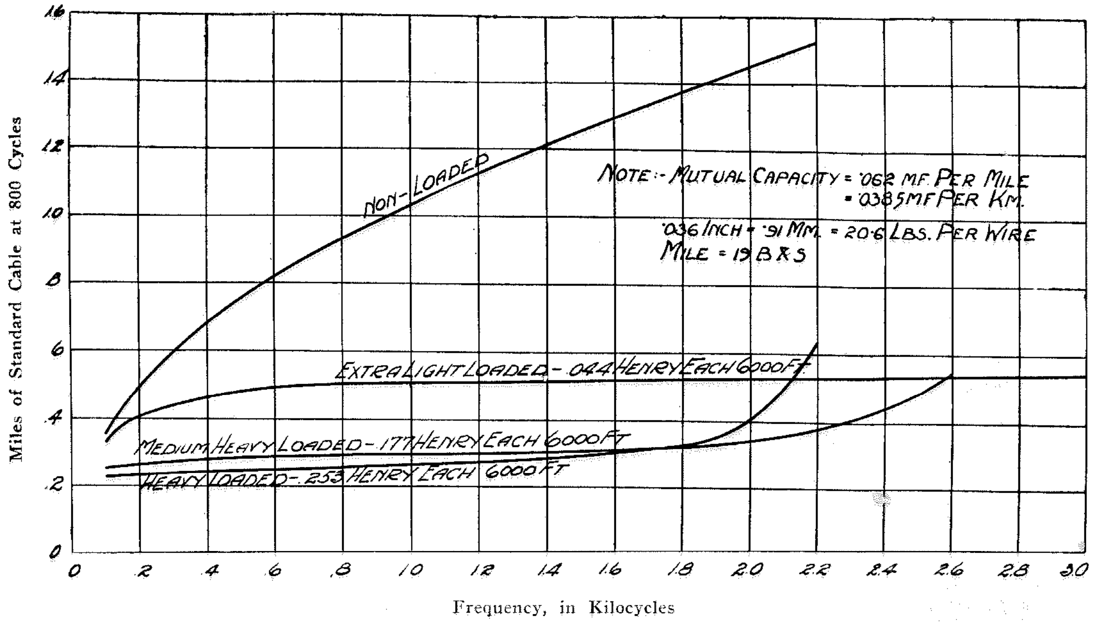


FIGURE 2

Figure 1 shows the loss for one mile of circuit at various frequencies in respect of open wire circuits weighing 435 pounds per mile (4.2 mm. diam.) both for non-loaded and for circuits loaded with 0.247 henries every 8 miles (12.9 kilometres); the computations are made for the steady state, that is, when the temporary effect of transients has passed off. From these curves we may see that the effect of loading has been threefold; (1) the attenuation has decreased, taking 800 cycles for example, from 0.035 to 0.016 miles of standard cable, a reduction in loss, that is an improvement in volume of speech, of 54%; (2) between about 400 and 2000 cycles the curve of loss has a greater slope for the loaded line, indicating that the various frequencies necessary to transmit satisfactory speech are less uniformly transmitted, thus increasing the frequency distortion and so degrading somewhat the quality which was of a very high order on the non-loaded line; (3) at about 2000 cycles, the attenuation of the loaded line undergoes a decided increase, termed the cut-off, so that high frequencies are extinguished. In addition to these three effects the speed of the circuit has fallen from

180,000 miles per second for the unloaded line to 55,000 miles per second for the loaded line.

Figure 2 shows the results of loading a circuit weighing 20.3 lbs. per mile in drycore cable with three different types of coil, each at a spacing of 6,000 feet (1829 metres). From this we notice four results, (1) the 800 cycle attenuation has fallen from 0.94 to 0.30 (taking the middle loaded curve as an example), a reduction of 68% in the loss, a greater reduction than was obtained in the case of open wire; (2) between 200 and 2000 cycles the loaded curve is approximately horizontal, indicating that all frequencies between those limits are almost equally transmitted so that the frequency distortion previously rather high, is made less, while the non-loaded curve has a pronounced slope, for example, the loss at 2000 cycles is 1.45, that is, 56% greater than at 800 cycles; (3) there is the same cut-off effect as was noticed on open wire lines when loaded; (4) the frequency of the cut-off point also falls as the loading increases. An additional effect is that the speed of the circuit decreases as the loading increases.

Effect of Loading Upon Cut-Off Frequency

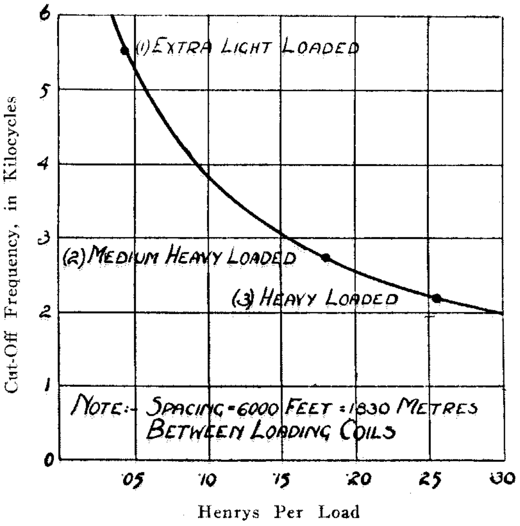


FIGURE 3

Effect of Loading Upon Velocity of Propagation

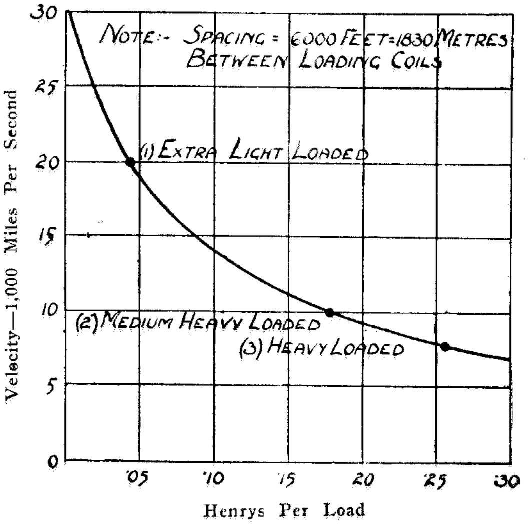


FIGURE 4

Figures 3 and 4 show the effect of loading on the cut-off frequency and upon the velocity of the cable circuits referred to in Figure 2.

It is seen that the addition of inductance to circuits, whether open wire or cable, affords a means of greatly extending the distance to which speech is possible, and of reducing the copper required to transmit speech over such distances as were previously feasible. But it is also seen that while loading reduces frequency distortion on the cable circuit, it increases this form of distortion on the open wire circuit; also that, by increasing the voltage in the circuit, it augments crosstalk and it reduces the speed of propagation in the circuit.

So far as open lines are concerned, the reactions on the general plant caused by loading are that a higher class of construction, including transposition and maintenance, is required to avoid crosstalk and to keep up the insulation, for loaded lines are much more susceptible to reduction in transmission efficiency due to lowered insulation than are non-loaded lines. With poor maintenance it may well be that the number of days in the year during which the improvement due to loading is gained is not sufficient to pay for the cost of loading.

This condition of constantly maintained high insulation applies also to loaded cable lines, but high insulation being comparatively easy to achieve in cables little reaction is caused by this. The increased crosstalk caused by loading has, however, caused real difficulty which has been overcome by great advances in the cable art, not only as regards the construction of long distance cable in the factory, but also as regards the jointing of the wires in the field, both these being intended to secure such freedom from unbalance between circuits as will obviate crosstalk. It must further be noted that here, as in many of the latest developments, the effect of variables is not necessarily local; that is to say, a defect in one place may be felt a long way off in a section where no defect exists. There was also at one time, but this does not now occur so often, a necessity for great care to guard loading coils from becoming magnetized; this has been avoided in cable loading coils by the use of compressed magnetic dust for the cores.

Repeaters

Although the telephone repeater has been in service since 1905, it is only since 1914 that the thermionic repeater (which followed the introduction of the grid or third element by de Forest into the two-element thermionic valve of Fleming), has been employed, and the great impetus to its use has only been given during the last five years or so. The fact that a three-electrode vacuum tube acts as an amplifier of speech currents has led to an idea that a telephone repeater begins and ends in a device which relays the received current very much as does an electromagnetic relay. But the telephone repeater has many important conditions besides the amplifying one and the fact that the reactions on telephone practice, due to the advent of the practical telephone repeater, have been and will be of very special importance, makes it worth while to devote some little time to their consideration.

First, let us look at the general types and their places in the system. There are so far, three general types: first, the repeater which operates in two directions by means of a single amplifying unit, the so-called 21-type repeater. This must be placed at or near the centre point of the line because the impedances of the lines on each side of the repeater act as balances to each other, and if they are not equal, the unbalance will cause circulating currents round the repeater with the result that sustained oscillations will be set up and the repeater will "sing." Up to the present, this type cannot be used in tandem, consequently, its use is limited; it may be applied either to open wire or to cable circuits. Second, there is the repeater which operates in two directions by means of two unidirectional amplifying units, the so-called 22-type repeater. With this type there is much greater freedom in locating the repeater because the balance is not between the two impedances offered by the lines on each side of the repeater, but between the impedance of one line and that of a network made to simulate the line impedance, and the precision with which the network does simulate its associated line at all speech frequencies, governs the degree of amplification or gain which may be taken from the repeater. If the balance is not held, circulating currents will cause the repeater to sing. Repeaters of this type are applicable to open wire and to cable circuits; they may be, and are regularly, placed in tandem, and as many as 23 have been used in tandem in regular service on a single conversation. This fact illustrates that the speech currents are transmitted with sufficient accuracy, as otherwise cumulative distortion would quickly cause degradation of articulation to an intolerable extent. Third, there is the repeater which operates in one direction only, the speech currents in the other direction being provided for by an independent circuit, the so-called 4-wire circuit, in which the currents from say, east to west, are taken by one circuit of two wires, with its unidirectional repeaters in tandem, and the speech currents in the other direction, from west to east, are taken by another circuit of two wires, also furnished with its unidirectional repeaters. Obviously, if a special 4-wire line were set up from Subscriber A to Subscriber B in which the circuit started with A's transmitter and terminated with B's receiver, and the circuit in the other direction were similarly treated, there could not be any circulating currents at all. Commercially, however, this is not possible, and it is necessary to use the regular 2-wire local system, so that only the long line portion of the circuit can be of the 4-wire type. With this type of repeater the circulating currents have to travel a long distance, which gives rise to great attenuation, before they can get back to their starting point for reamplification, and so a much greater gain can be taken from 4-wire repeaters than from 21 or 22 types, before the singing condition is approached. While the 4-wire type can be employed on either open or cable circuits, the fact that it requires 4 wires makes it economically more suited to cable circuits, and since the gains obtainable are high (they can be made so high as to render the line loss zero between the terminals of the 4-wire section, which cannot be achieved by any other type), it is economically possible to employ this type of circuit in cable for distances up to 1,000 miles, perhaps further, and so it is preeminently suited for groups of long distance lines carrying heavy traffic.

Gain-Frequency Curves of 22 Type Repeater

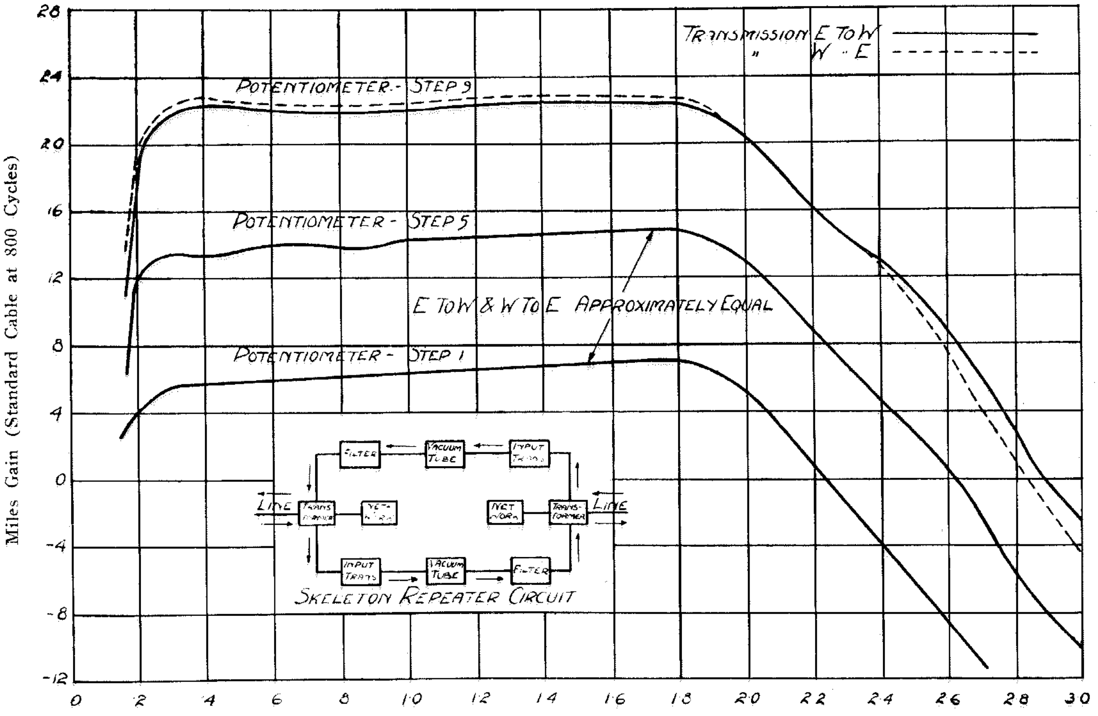


FIGURE 5

Whichever repeaters are employed in any line, they must, of course, be located at the right positions determined by engineering considerations alone and not by political ones. It might, for example, be correct for a line to run through Switzerland with no repeater on it at all in that country, or for a line to run through Limburg on which there would be a repeater in Holland, but no appreciable length of line in that State. In any such case the networks at the repeater in one country must conform to the lines situated in other countries.

The fact that telephone repeaters must be employed in tandem renders their requirements very severe. Figure 5 illustrates the gains obtained from a repeater of 22 type designed for a loaded cable circuit taken at random and connected to an artificial line. The control of the gain is by a potentiometer with fixed steps, the top curve showing the gains from the two repeaters in the two directions, east and west, for speech frequencies when the two potentiometers were set on the same step, the 9th. The gain shown in the curve is approximately 22 miles, a current amplification of 11 times. Of course it will be realized that this only shows what amplification the repeater can give, in practice one would not expect such gains when connected to real lines. It will be seen how closely alike are the gains in each direction; that is, the speech currents are amplified in either direction with very nearly identical gains. When the potentiometers were set on the 5th step the gains were reduced, but the curves of the gain in each direction were practically indistinguishable, and the same held good when a still lower gain was taken by puting the potentiometers on the 1st step.

It will be noticed that these results were obtained not by any careful and fractional adjustment, but merely by setting the 2 potentiometers on similar steps, and it will also be seen that the variation in gain produced by one step is approximately two standard miles. The small inset diagram indicates the pieces of apparatus involved, all of which have to play their part in the gain at speech frequencies, viz., potentiometer, input transformer, vacuum tube, output transformer, filter and three winding transformers in each repeating unit. It will be realized that the gain given out by the repeater must, of course, be adjusted to and suitable for the type of line with which it is to be used; also that a repeater is not a universal article which can be attached to any line, regardless of its make-up.

Impedance-Frequency Curve of .118 Inch Open Wire Non-Loaded Line

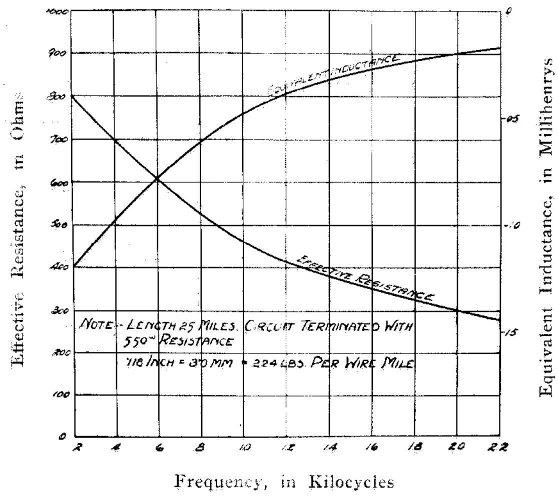


FIGURE 6

Impedance-Frequency Curve of .157 Inch Non-Loaded Open Wire Side Circuit

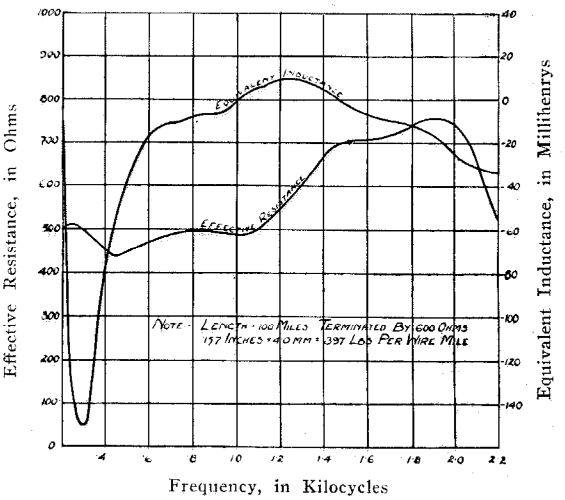


FIGURE 7

When in actual service repeaters are associated with lines, several reactions occur which very largely modify previous practice. As before stated, the line is balanced by an impedance network which simulates the line, but in order to keep these networks practicable it is necessary that the lines shall be as free as possible from irregularities, otherwise the networks would be very expensive and perhaps impossible. Figure 6 shows impedance frequency curves for an open wire line 41 km. (25.5 miles) long of copper wires having a diameter of 3 mm. (0.108 in.). This is an excellent example and so regular that there is no difficulty in providing a network which closely simulates the line, consequently satisfactory repeater gains can be obtained. The indicated results were found at the first trial and without any clearing up. Figure 7, however, shows similar curves for a line that was regarded as a first-class one until tests were made on it prior to using a repeater. It is seen that its impedance curves are very irregular so that in its then state it was quite impossible to employ a repeater. Figure 8 shows a cable circuit 34 miles (54.7 km.) long, from which, for some unknown reason, a loading coil 8 miles (12.9 km.) distant from the place of this test, had been removed. The

Impedance-Frequency Curve .066 Inch Light Loaded Cable

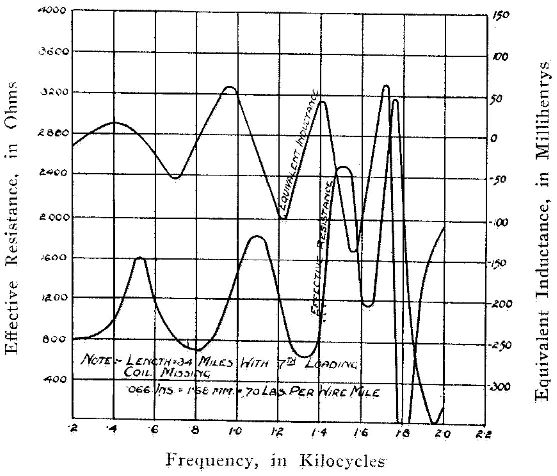
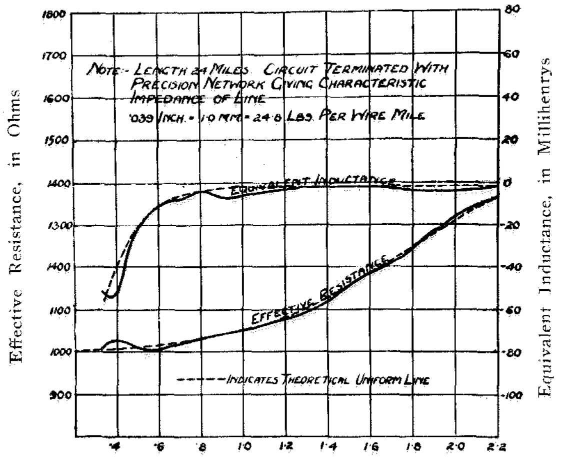


FIGURE 8

impedance frequency curves very plainly show how, by this removal, the line is thrown out of balance with the network, and one can thus see that the removal of the coil would at once render useless the line of which it forms a part. The removal of a loading coil is a noticeable matter, but the want of uniformity introduced by several portions of a line having different constants, such as non-uniformity caused by the haphazard joining up of lines not constructed with a view to the rigid uniformity required for repeatered lines, acts in the same fashion.

Figure 9 shows impedance frequency curves for a phantom loaded cable circuit 39 km. (24 miles) long. It will be remembered that a phantom circuit is one which is obtained by superimposing the phantom on two circuits each composed of two physical wires. It is not an easy matter to obtain a low unbalance condition between the various capacities which make up such a circuit. In this figure are shown the curves for the line and for a theoretically uniform line; it will be seen that the actual line does closely approximate to the theoretically perfect line.

Impedance-Frequency Curve .039 Inch Phantom Cable Circuit
Medium Heavy Loaded Cable



Frequency, in Kilocycles

FIGURE 9

Let us now see what is the overall result in transmission when employing a long line in which copper and repeaters both contribute to the effective transmission of speech. What, in fact, is the transmission afforded over the whole system? In Figure 10 are shown frequency attenuation curves for a non-loaded open copper wire line 3,400 miles (5,472 km.) long. The curves reproduced are, A, actual measurements on the line when using repeaters of a type designed for and suited to the line; B, computations based on using imaginary repeaters giving uniform gains at all frequencies; C, actual measurements when using repeaters suitable for other types of line but unsuited to this one. The line included 12 repeaters in tandem and the results are somewhat remarkable. They show, curve A, that the specially designed repeater in conjunction with the suitable line gives a fairly uniform overall loss, approximately 10 standard miles, between the frequencies of 400 and 1800. On the other hand, curve B shows that the theoretical uniform-gain repeater, if used, would be very unsatisfactory, in that it would give a frequency gain characteristic of a very undesirable kind and one which would greatly increase the frequency distortion. Lastly, curve C shows that a repeater suited to the character of one line, if used with another line to which it is unsuited, may give overall transmission of a highly unsatisfactory nature.

Since it is not yet practicable to transmit all frequencies equally, it is evident that some sort of a compromise must be made, and if a line is composed of several sections on each of which a different compromise has been made, the final through result may be less satisfactory than need be solely because of the fact that the compromises contain no unity of treatment.

When repeaters are in operation, they must maintain constant the gains to be given out or there will be serious effect on the speech. If we assume a 4-wire circuit between Rotterdam and Milan, 500 miles (810 km.) long and having five repeaters in it, operating at gains of 23, 30, 30, 30 and 23 standard miles (S. M.) respectively, say an average of 27.2 each, we need only consider what will happen if the gains fall off, since the gains will originally have been set to be as high as safely allowable. Assume, then, that the line without repeaters has a net equivalent of 148 S.M. from which we deduct the repeater gains, 5 x 27.2 = 136, leaving the net loss = 12 S.M. Now suppose the gain at each repeater station for any reason at all falls off by 2%, this will represent 0.54 S.M. each or 2.7 S.M. for the five stations, and the net result will then be increased from 12 to 14.7 S.M., an

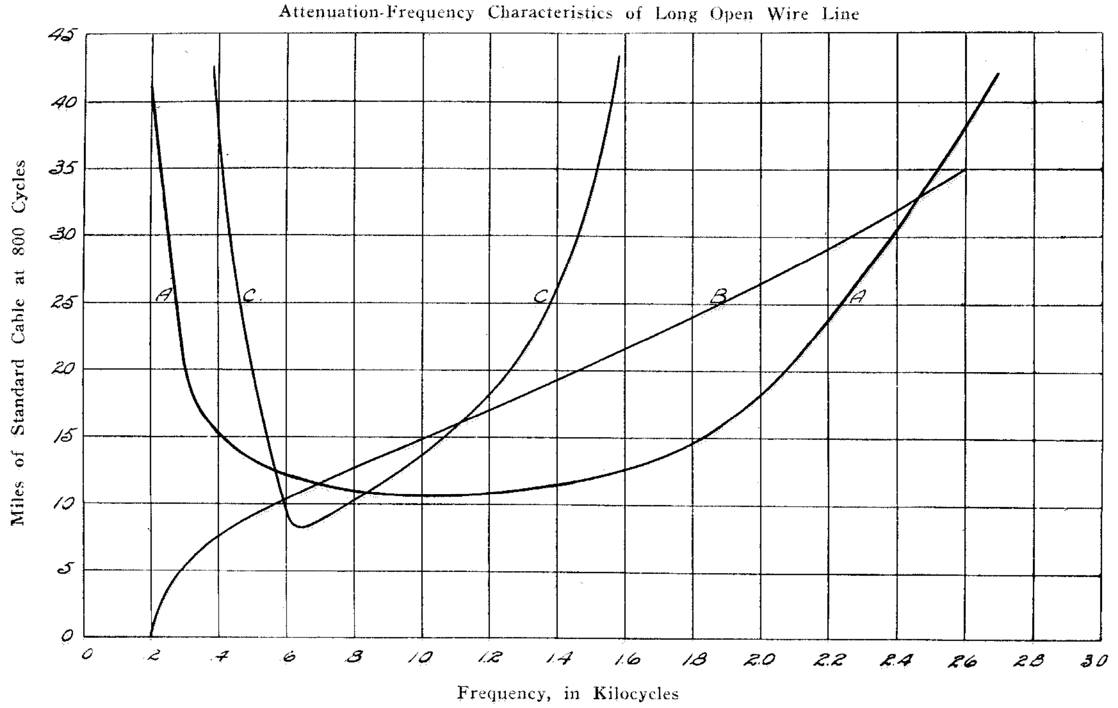


FIGURE 10

increase of 23% in the loss in the line. Should the gain on each repeater fall by 7.5% the total additional loss will be 10.1 S.M., and the final net loss will be increased from 12 to 22.1 S.M., an increase of 84.5%. In this case the loss would be so great that probably the line would become unworkable. I have chosen these examples to show the importance of uniformity of construction, uniformity of maintenance, and uniformity of operation; it will be seen afterwards what is their particular application. The examples are rather understated than exaggerated. It would have been quite reasonable to have taken a case with twenty repeater stations in tandem, and, furthermore, the gain given by a repeater would not in fact be one definite figure for all frequencies.

Fortunately, the design of repeaters has been carried far enough so that if correct design is employed and if certain regulations for operating routine and maintenance are followed, the gains can be held steadily, but among those routines are tests which determine when the useful life of an amplifying element, a vacuum tube, has ceased and the required constancy in gain can only be held if all repeater stations are operating to the same routine. If the line is an aerial one and subject to considerable changes in temperature, the resistance alters and another source of variation in overall transmission equivalent is introduced. If these changes are serious they can be compensated automatically, if not so serious, they can be dealt with by operative routine. The lesson is, however, the same in either case, and, for the best results, those persons operating the various repeater stations must be operating to the same routine, employing the same technique and under the one control. Again, when breakdowns occur important circuits cannot stand out of order, but must immediately be temporarily rerouted to restore the service. Such alteration may affect the balance between the network and line at the repeater station and this may have to be dealt with by altering the gain of one repeater (in which case alterations will probably be required at all other repeater stations along the line), or by changing networks. If it were possible to foresee all possible combinations and emergencies, it would doubtless be possible, though not economical, to establish routines covering all cases requiring attention, but obviously this is not possible and the only commercial solution lies in unity of control of the line from beginning to end.

It has been shown that the result of loading the circuit is reduced attenuation and somewhat impaired articulation in open wire circuits, and reduced attenuation and better articulation in cable circuits. Now that the use of repeaters has become possible, additional energy can be put into the line as required and the attenuation can be reduced by that means. It is therefore no longer necessary to sacrifice the quality which can be obtained on open wire circuits by loading them in order to reduce the attenuation; this reduction can be effected by repeaters. In cable circuits, however, it was shown that loading was necessary to reduce the frequency distortion. Consequently, long, heavy open wire lines are not now loaded at all but are repeatered, resulting in improved articulation, and the increased speed of propagation avoids echo trouble which only became insistent because of the more powerful effects derived from repeaters. With cable circuits, on the other hand, loading still obtains, it cannot be abandoned since it is necessary for the reduction of frequency distortion, but the tendency is towards lighter loading so as to raise the speed of the circuit, thus reducing the echo trouble which, because of the reduced speed and the great electrical length of loaded cable circuits, demands most careful consideration.

CARRIER CIRCUITS

In the search after increased capacity of telephone and telegraph circuits, there has recently been developed and put into commercial service the carrier system which has been added to the well-known methods of superimposing phantom telephone and compositing telegraph circuits. In this new method carrier waves of different frequencies for each channel of communication are generated. If the channels are to be used for telephony such waves have a frequency above the audible limit; by means of band filters the desired range of frequency is permitted to pass into each channel but only frequencies within that range. Thus on a 4-channel telephone carrier circuit the frequencies might range in four or eight separate bands with outside limits of 4,000 to 27,000 cycles per second. Each carrier wave is modulated independently by the voice currents to be transmitted by that channel, and all the modulated carrier waves, or all of one of the side bands only, without the carrier waves, are transmitted over the line. Upon reaching the far end, the waves are filtered out each into its proper channel according to the carrier frequency assigned to each channel, and are then demodulated leaving the voice current free to be farther transmitted over an ordinary circuit. Because of the increased frequency of the carrier waves, greater attenuation occurs with them than with the voice waves and carrier current repeaters must be equipped more frequently than voice current repeaters. Also, for the same reason, carrier currents cannot be transmitted over ordinary loaded lines, which, it will be remembered, cut off at frequencies within the audible range. Hence if loaded carrier circuits are required they must be specially treated. Special treatment is also needed in the construction and maintenance of carrier lines and equipment, and because the equipment is expensive such lines must be of considerable length in order to be economical.

As an illustration of the advantages to be gained by using the latest development, the following may be quoted. On the New York-San Francisco line, the circuits are of open wire from Harrisburgh to San Francisco, about 2500 miles (4050 km.) direct distance apart. On four conductors on this route the loads carried are:

2 Physical Telephone Circuits
1 Phantom Telephone Circuit
4 Earthed Telegraph Circuits, and
a varying number of Carrier Telegraph Circuits, ranging from 6 to 20.

Two of the sections on this route in detail are:

Between Chicago and Omaha, 450 miles (729 km.) direct distance apart, 4 open wire conductors carry:

2 Physical Telephone Circuits
1 Phantom Telephone Circuit
4 Earthed Telegraph Circuits which can be worked either 1 way or 2 ways at will.
20 Two-way Carrier Telegraph Circuits
27 Total Circuits on 4 wires.

Between Chicago and Pittsburgh, 450 miles (729 km.) apart, eight open wire conductors carry:

4 Physical Telephone Circuits
4 Halves of Phantom Telephone Circuits (equivalent to 2 circuits)
8 Earthed Telegraph Circuits which can be worked either 1 way or 2 ways at will
37 Two-way Carrier Telegraph Circuits
51 Total Circuits on 8 Wires.

From another route we take the following:

Between New York and Philadelphia, 90 miles (145 km.) apart, two conductors in cable carry, though not by carrier circuits:

1 Physical Telephone Circuit
30 Special Signaling Circuits
31 Total Circuits on 2 wires

From Chicago to Omaha or from Chicago to Pittsburg, the direct distances are about the same as from Paris to Berlin, from Paris to Marseilles, or from London to Milan. From New York to Philadelphia is about the same distance as between London and Birmingham.

At present there are in actual service in the United States the following miles of carrier route and channel:

	Miles of Route	*Miles of Channel*
Carrier Telephone	4-776	16 576
Carrier Telegraph	10,919	78,870
Total	15,695	95,446

CROSSTALK

Applications such as have been described demand a much higher degree of refinement in order to avoid crosstalk, than those which have previously obtained in the construction and maintenance of long distance lines. To obviate that evil, it is necessary that at every point throughout the entire length telephone lines should have the two sides of the circuit equal in admittance to earth and equal in series impedance, and these must be equal over the range of voice frequencies. This is a very severe requirement, but very good approximations to the result required are being made.

INTERFERENCE

A matter which is assuming more and more importance is that which in the communication art is termed interference, meaning by that term the reactions which occur between weak current communication circuits and heavy current light power and traction circuits.

The effects of these reactions to the communication engineer may be serious and fall under the heads of:

Noise
False Signals
Breakdown of the Line
Fire Hazard
Acoustic Shock
Electric Shock

Some consideration has already been given to the question of balancing the telephone circuits, and before looking at the same matter in regard to the power lines perhaps it may be useful to give an idea of the relative trouble caused by different frequencies.

Relative Interfering Effect of Single Frequency Currents in a Telephone Receiver

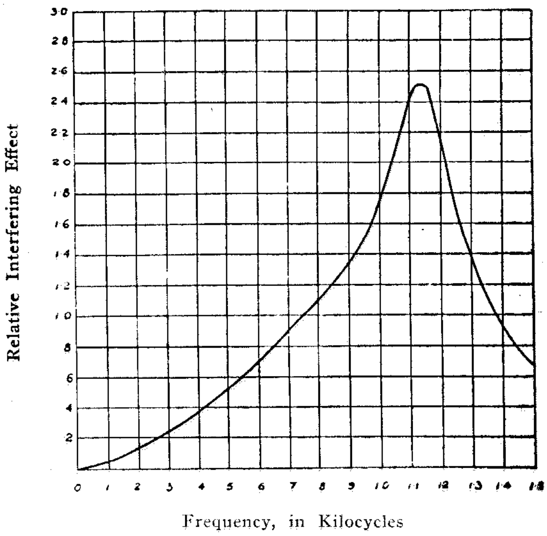


FIGURE 11

Figure 11 shows the relative interfering effect of uniform currents at various single frequencies in a telephone receiver; the interfering effect is very unequal and the importance of the wave shape in power circuits will be inferred from this curve.

On the power side, residual and balanced components of the power circuit voltages and currents may cause such trouble as to be beyond the ability of the communication engineer to cure. Every commercial 3-phase system, for example, which has not been properly transposed, is an unbalanced system, and any change in the separation of wires or height from the ground will affect the balance to earth which generally is of more importance than balance of load between phases. This unbalance can be much reduced by transposing the power lines. Again, even if well balanced during normal operation, power lines are invariably thrown badly out of balance by abnormal occurrences such as the opening or short-circuiting of the line; and sometimes the circuit and switching arrangements are such as needlessly cause unbalanced effects not perhaps noticed by the power engineer, but very troublesome, if not worse, to the communication engineer.

Now it must be recognized that these industries, involving the telegraph, telephone and railway signaling systems as representative of the light energy group and the lighting, power, railway and tramway systems as representing the heavy energy group—are both of them necessary to the well-being of the world, and they must learn to live together harmoniously, and to avoid or mitigate the otherwise serious reactions between their respective circuits. It must also be recognized that in grappling with this difficult problem there cannot and ought not to be any claim by either side for priority of protection or preferential treatment. It is wrong for the heavy current interest to say—let the light current industry take care of itself—and it is equally wrong for the light current industry to say—the heavy current business must be conducted in such a manner that we, with our existing arrangements, shall be undistrubed. There is only one sensible solution: let the engineers of the two industries first get together, unfettered by any partisan tie, to seek the best methods of getting rid of the trouble, and after those best methods have been found, on the basis of the least total cost, then, and only then, let the question of settling the apportionment of cost as between the interests be taken up. The ordinary difficulties of a complex situation are frequently rendered more difficult of solution by an endeavor at the outset to fix responsibility for the interference.

Much has already been done by joint study to reduce interference and in some cases, such as those of electrolysis, it has been found economical to the heavy industry to avoid certain defects in construction which were first brought to light by the complaining light industry. But in all cases the lesson is always being pressed home that success is certain to come when each party makes a real endeavor to learn the other's problems and to appreciate the efforts made by him to solve them.

Heretofore a long distance telephone line was a relatively simple structure consisting merely of a pair of copper wires, either open or in cable. This could be maintained comparatively easily in good order by independent maintenance units situated along the length of the line. But with repeaters and loading that simple structure has vanished, the plant is more complicated and the various parts are interdependent on each other. It is no longer possible to consider maintenance of each part solely as a sectional matter; what is done at one place may cause serious reactions at another, and the line as a whole must be considered.

We can now, therefore, obtain certain advantages in the construction of through lines, but only if we are willing to give the attention necessary to secure them. It is false to imagine that we can obtain the benefits of the present knowledge without taking the necessary steps to secure them. The benefits are:

Great increase in the distance over which communication can be given.

Great increase in the number of channels of communication, telephone and telegraph, which can be provided by one pair of wires.

Great increase in the number of circuits which can be placed in cable, numbers such that it would be impossible to find space for them if all circuits were to be open wires.

Greatly reduced annual cost of circuits and improvement in quality of speech.

Increased security of service by reason of circuits being in cable.

Increased speed of service by reason of greater number of circuits.

The principal points to which attention must be given to secure the above advantages are:

Definite decision as to the work each line is to do, that is, planning in advance.

Definite standards of performance to be required of the complete line.

Unity of treatment of all transmission matters affecting the line over its full length.

Unity of treatment of all transmission matters affected by the connecting of the line to other lines, whether trunks or subscribers' lines.

Unity of maintenance control over complete length of the line.

Unity of control over repeater gains over the complete length of the line.

Unity of operating control over the whole length of the line.

Education of all sections of the telephone staff in transmission, maintenance and operating practices. Without this education among all detail members of the staff it is impossible to obtain the benefits now available. It is not sufficient for a few engineers in any administration to be familiar with these matters; they must be made part of the general knowledge of all, and this is particularly true of education in matters affecting transmission which must by some means or other, in the varying degree required, be made to permeate all classes of the staff who have to do with the transmission plant.

Greater attention must be given to the economic aspect of engineering studies for the question of economics is the fundamental problem of the engineer; let him neglect it and even though brilliant, he becomes something else, a physicist perhaps, a mechanician, or a constructor but not an engineer, who, it seems to me, must ever link together the progress of science with the minimum of human effort, that is with the most economical manner of achieving what is desired. In the construction of plant how often it happens that when the engineer has decided what plant shall be installed, the question as to whether the plant shall or shall not be profitable is also decided; for when the engineer has constructed the plant and gone elsewhere, the management can only affect results by a margin, important no doubt, but the fact remains that the decision as to what plant shall be installed settles also many other items of expense, both for labor and material.

The telephone service has certain special features, in that (1) the unskilled public is an actual participant in a call, the matter is not merely handed over to a skilled operator although the skilled operator also participates; (2) it is essentially a through service, *i.e.*, the whole circuit embracing the calling and the called persons' instruments and the complete line connecting them are simultaneously in use for each call. Therefore, all parts of the circuit must be harmonious, although these parts may belong to and be operated by different owners. (3) The operators at the various stages along the line have to cooperate with each other and differences in the operators' technique will decrease efficiency. (4) It is a world service, for it is impossible to set any limits to the service which must extend as the degree of technical knowledge permits.

Frequently in industry one cannot obtain an absolute standard and recourse must be made to relative comparisons. It is so in telephony, and as I assume I may take it for granted, at any rate by those who have studied the matter, that the telephone systems of the United States are in advance of those operating in Europe, it is worth while to see wherein lie the differences (altogether apart from ownership) and particularly the differences in organization, and to obtain some idea of the telephone system in the United States, which now has nearly two-thirds of the telephones throughout the whole world. In that country there are at present over 10,000 companies owning and operating over 14,000,000 telephone stations—that total number divides into two broad classes—those having some kind of connection with the Bell System and those which have not. Again, the first class divides into those known as Bell-owned, and others as Bell-connecting with an independent ownership.

In his presidential address in 1905 Sir John Gavey referred to the growth of the Bell System as "absolutely startling." He said there had been an increase of 1,450,000 stations in seven years—an average of 207,000 per year. But since then the increase in the Bell owned Companies has been an average of 410,000 per annum, or twice the number which startled Gavey.

We may tabulate the telephone statistics of the United States thus:

		July 31, 1922	
	No. of Companies	*No. of Stations*	*Per Cent.*
Bell owned Companies	26	9,223,770	65.0
Bell connecting Companies (Independent ownership)	9289	4,520,725	31.8
Total Bell System	9315	13,744,495	96.8
Non-Bell-connecting Companies	879	452,597	3.2
Total	10194	14,197,092	100.

Taking the population of the United States at one hundred and nine millions it will be seen that there is now one telephone station to 7.7 persons, while in the year 1900 there was only one telephone station to 56 persons. Since the beginning of the twentieth century while the population has increased by 45% and the volume of general business (judged by the best data available) has increased by 100%, in the same time the number of telephone stations has increased by over 900%.

Again, if we judge progress by capital expenditure we find that the investment of the Bell owned Companies which was $180,700,000 in 1900 had increased by 267% by 1911 and by 755% by 1921, and then stood at a total of $1,543,865,545—say £346,000,000.

As a method of trying to give an impression of the telephone service in the United States it may be said that from his telephone in that country a subscriber can reach out over more than 4,000 miles and can call practically any of the 13,700,000 stations referred to situated in 70,000 cities, towns and villages, and the statistics show that the telephone communications in that country outnumber the postal communications by fifty percent. It is agreed by those best qualified to judge that American industry on its present scale could not function without the telephone service as they know it there.

From these figures it will be seen that while there are many telephones which are not part of the Bell System, yet the great majority (97%) are part of that organization of companies, and further it may be stated that with a few exceptions those companies which are not part of that system are on the average a collection of small concerns. In what follows and generally in connection with the expression "telephony in the United States," the Bell System is referred to.

There are five outstanding features in the organization of the Bell System, and I think it may be said that these features are essential in any effective organization for telephony on any extended scale.

The five features are:

1. Local operating organizations thus making for decentralization. These organizations, or companies, possess large measures of authority.
2. A central administrative direction and control over the local organizations.
3. A long distance organization constructing and operating the long lines by which the local organizations effect inter-communication.
4. Control of the manufacturing organization.
5. A central organization for scientific research, development of apparatus and technique of construction, maintenance and operation.

In Europe, generally speaking, and considering the nations separately, we find:

1. An organization having a central authority with no separate local authorities.
2. A series of administrative areas charged with the duty of maintaining the service under the central authority.
3. No one department charged with the duty of through business.
4. No control over manufacture.

When we consider Europe as a whole we find:

1. A number, about 40, self contained local operating organizations each, in the majority of cases, conducting a local business and a through business within its area; also that part of the international through business which lies within its own borders.
2. No organization controlling or coordinating the various local operating organizations which yet have to function as a whole.
3. No means of keeping the separate organizations in touch with each other, and no systematic means of adjusting differences in matters of daily practices.
4. No organization of any kind which handles and cares for the through business as a whole.
5. No common agreement as to manufacture.
6. No common research, standard practice or technique of construction, maintenance and operation.

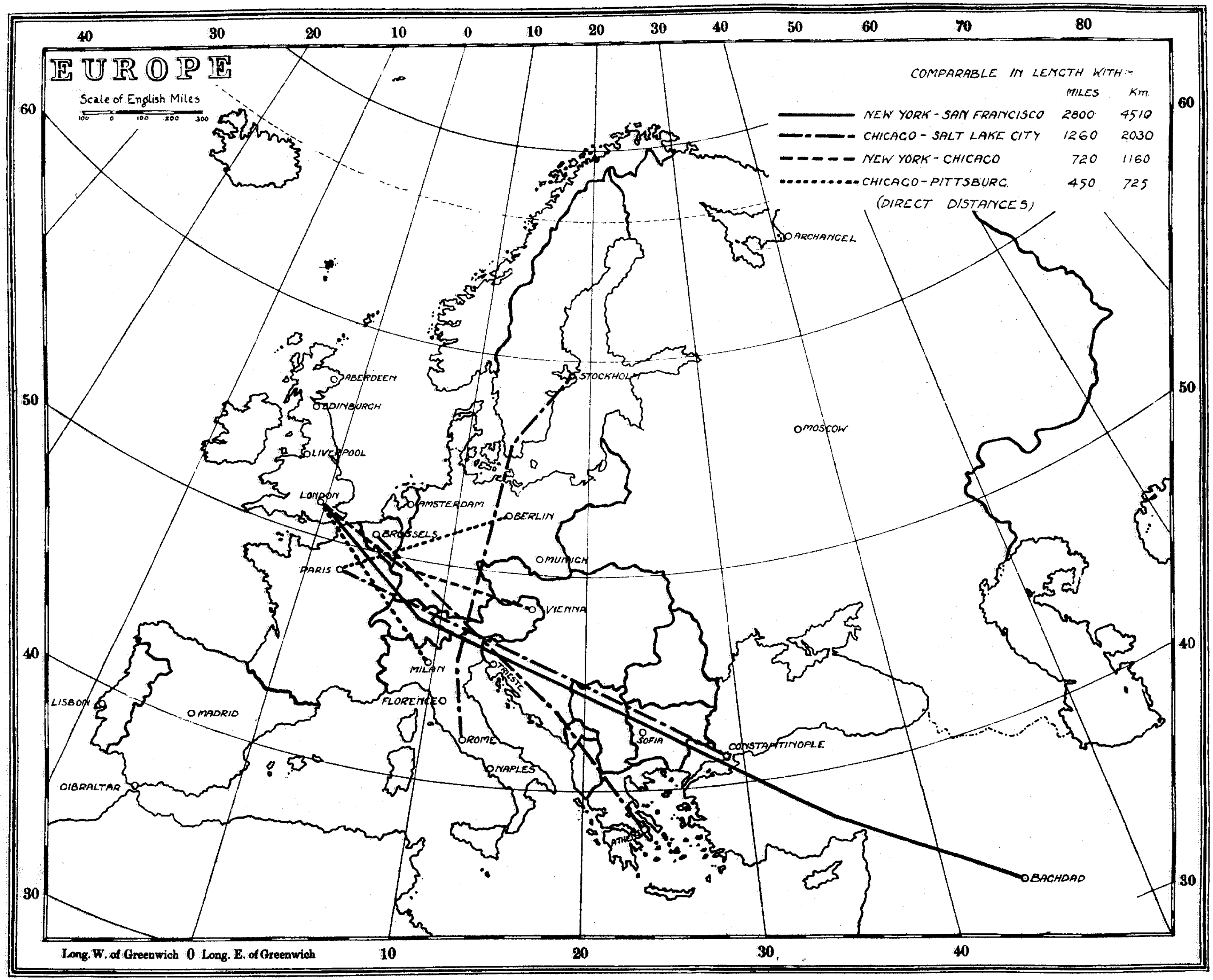
At the moment we are not concerned with the effect of this loose coupling upon the local business of each country, but little consideration is needed to appreciate its harmful effects upon the through business between countries whether the length of line over which such business is conducted is great or small. There are in Europe large centres of population within such distances of and in such commercial relationships to each other that traffic would be forthcoming did adequate facilities but exist. There is no engineering difficulty so far as distance is concerned, in constructing and operating lines at commercial rates to give satisfactory speech from any part to any other part of Europe, but at present the through business is meagre in quantity, slow and inefficient. Under the present conditions, practically the only way in which the nations can cooperate in these matters is that when new lines are to be constructed between countries there is cooperation and consultation between the representatives of the countries concerned and occasionally there are international conferences. But these do not, and cannot, produce a unified system; all that they can do at the best is spasmodically and partially to compromise on a few outstanding differences in practice, which between whiles grow up unchecked, and to leave unsettled such large questions as cannot be agreed.

The settling of arrangements, and particularly the financial arrangements, for the construction of additional direct lines between contiguous countries constitutes an operation difficult enough, but when it is sought to construct lines between non-contiguous countries, in which cases they have to traverse countries which are not interested in the traffic desired by the terminal countries, the difficulties in the way of getting anything done are great indeed, and much praise is due to the energy and enterprise of those men who have succeeded in achieving the service now in operation.

Yet there is every indication that given facilities there is traffic waiting to be handled between the cities of Europe as between the cities of the United States. The opinion of some of those well qualified to judge is that the differences in language and customs do not, as they would at first sight appear to do, constitute a serious bar to international communication by telephony, and there are weighty reasons such as the present necessity of improving the relationship between nations, in addition to the normal commercial advantages which render it safe to forecast sufficient through business to warrant the setting up of a competent organization with the plant necessary to handle the traffic.

But there is little likelihood of speedy and economical construction and operation of such lines as are necessary between, say, London and Stockholm, involving three or perhaps five intermediate non-interested countries; London and Christiania, involving perhaps six intermediate countries; or London and Petrograd, involving eight intermediate countries, and yet there is nothing fanciful in the idea of quick communication between such places. The direct distance between Brussels and Athens, or Paris and Constantinople, is 1,300 miles—about the same distance as between New York and Omaha, or between Chicago and Salt Lake City, between which places calls can, at any time, be made. The direct distance over land between London and Bagdad is about the same as between New York and San Francisco, over which line conversations take place daily, while the direct distance over land between London and Delhi is about the direct distance from Key West in Florida to New York, thence to San Francisco and thence to Los Angeles, in California, over which distance calls can be made regularly. As a further encouragement, it may be said that the New York-Chicago cable now in course of construction, will have a gross transmission equivalent so great that if a 435 lb. (4.2 mm.) open wire circuit were constructed to that equivalent it might be 10,000 miles long, enough to connect Paris to the telephone system at Seattle in the north western part of the United States and leave enough to spare to take care of the cable across the Bering Strait. Of course, this illustration is an uncommercial one, but it serves to show that land distance is now no difficulty to telephony.

If we consider such business in the United States, we find that there are originated at New York over 4,000,000 long distance calls per annum, and it will be remembered that in the United States many calls are made over lines of considerable length belonging to the Local Companies, and do not go over the long distance lines. Similarly we find that Chicago


EUROPE
Scale of English Miles
100 0 100 200 300
COMPARABLE IN LENGTH WITH:-
MILES Km.
NEW YORK - SAN FRANCISCO 2800 4510
CHICAGO - SALT LAKE CITY 1260 2030
NEW YORK - CHICAGO 720 1160
CHICAGO - PITTSBURG 450 725
(DIRECT DISTANCES)
ARCHANGEL
STOCKHOLM
MOSCOW
ABERDEEN
EDINBURGH
LIVERPOOL
LONDON
AMSTERDAM
BERLIN
BRUSSELS
PARIS
MUNICH
VIENNA
MILAN
FLORENCE
TRIESTE
ROME
NAPLES
SOFIA
CONSTANTINOPLE
ATHENS
BAGHDAD
LISBON
MADRID
GIBRALTAR
40 30 20 10 0 10 20 30 40 50 60 70 80
60 50 40 30
Long. W. of Greenwich 0 Long. E. of Greenwich
10 20 30 40

and Philadelphia each originate a number approaching 2,000,000 long distance calls per annum, while such places as Boston, Cleveland and Pittsburgh each originate about 500,000 long distance calls per annum.

Figure 12 is a map of Europe on which are shown a few of the long distance circuits in the United States which are in regular daily commercial operation. It would have been easy to show a great many more, but this is not necessary in order to bring home the simple fact that there exists in that country a long distance telephone service such as is not known in Europe. And yet there is no reason whatever why the service in Europe should not be extended in a somewhat similar fashion. From the fact of its denser population and less distant cities, Europe enjoys advantages over the United States, and these should make for much greater development of the through business than she now has.

It is not putting the matter too strongly to say, that through telephony in Europe under the present conditions can never be worth the name of a service, and that the alternatives are either forever to be condemned to an ineffective, inefficient state of affairs, or to find some plan, other than the present one, for dealing with the through business.

Analyzing the conditions of through telephony in Europe as a whole, it is obvious that each nation, sovereign though it may be within its territory, is really, from the telephone point of view, merely conducting a local business over an area which is not very great. It is also clear that no one local authority can operate its own through business outside its own boundaries; although vitally interested, it must at its boundaries hand over the conduct of its business, in part, to someone else.

The through business must be handled as a complete unit if it is to be efficiently done, it cannot be done by independent units. The examples of recent improvements which I have referred to have been selected mainly because they illustrate the unity of treatment required by long lines. The correct course, therefore, appears obvious, viz., to depute a body to do for all European nations that which no one nation can do for itself; this is not a new departure, it is already practiced by banks and railways in their clearing houses. No bank would now tolerate for a moment any attempt to effect for itself clearance of the various cheques presented to it daily, and consequently we find that the banks themselves have established their clearing house, which performs specialised functions for all banks, and thus expedites the work of all.

The corporate spheres assigned to any telephone authority may be determined by political, financial, legal or other considerations, and by reason of these spheres and considerations, the authority is entitled to receive revenues and is obligated to pay the taxes and bills arising out of or payable in respect of those spheres. But these corporate spheres have, in reality, nothing whatever to do with the operating areas, which ought to be fixed solely with regard to obtaining the most efficient operating possible, and without any regard whatever to the corporate spheres. If the two differ, it is quite feasible for the operating authority to account as between any two or more corporate spheres without sacrificing any operating efficiency. Once the fact has been grasped that there is no reason whatever for the corporate spheres of influence and the operating areas to be identical, and that each requires quite separate consideration for its determination, there will be no real difficulty in arranging operating areas for efficiency and apart from corporate spheres.

With sectional, non-unified control over the various portions of the through business, it is not possible to design, construct and operate through lines of communication in a manner capable of meeting the needs of the public. It has already been shown in what manner conditions in one part of the plant may react on conditions at another part, and how these parts may be distant from each other, so that in fact what is done in one country may render ineffective the efforts made in another country. It ought not to be necessary to labor this point, but perhaps an analogy may help. The through business is as much a unity as is military operation. We have seen the advantage gained by unity of command in warfare and no one would now advocate independent multi-commands such as were seen in 1914 and the early years of the War. If it were possible for the nations to agree on such unity of control for the purposes of War, it ought not to be beyond their powers to agree to a unity of control for the efficient working of the through telephone business. It is not enough for the separate organizations to attempt to agree to a code of rules to which each shall subscribe—such an attempt would only be to court failure. The business is varying, flexible and very much a living thing, it demands intelligent and prompt treatment of its many variations, it requires control from central points carrying with it the power to instruct persons at great distance in the routines and duties they are to perform and such control can only be effected by a living authority always on duty.

Besides the engineering considerations which have been dealt with, there are weighty reasons connected with the matters of circuit layout, business policy, rates and operating, about which much might be said, showing the impossibility of giving an adequate through service without unity of control, but this is not the place to deal with them.

It is easier to analyze the conditions and to state the fundamental requirements for efficiency, than it is to propound a scheme for an effective organization. Yet some effort at a solution must be attempted even though it is unlikely that the first attempt will be successful. Any solution must find some method of satisfying the financial needs of the business as well as the technical requirements. At present it is difficult enough for the various administrations to obtain from their Governments the money required for the construction of such plant as is demanded by their own traffic, let alone for the fostering of traffic by the construction of lines not yet called for by public demand, and for the construction of lines between non-contiguous countries, which lines although demanded, are not required by the natives of the intermediate countries through which they pass. In fact, in spite of the reality that Governments can borrow money at a cheaper rate than can public companies, it remains true that Governments do find difficulty in raising the capital necessary for the legitimate demands of telephone development.

The alternative suggestions which present themselves are:

1. To operate all the through business both within and between the various countries in Europe by a single Long Lines Company working under licenses from the various Governments, taking the calls from the local originating organizations, and being entirely responsible for them until turned over to the local receiving organization.

Governments would put the Long Lines Company into a proper legal position, and make it plain that the Company had the goodwill and support of the country, and they would co-operate with the Company in the handling of the traffic. It might also be found desirable to turn over to the Company, either on purchase or rental, certain lines and equipment already in existence for handling through traffic.

The advantages of this course would be that unified control could be achieved at once. The service would be on an ordinary commercial basis and if the fees were correct, sufficient money could be raised to construct all lines and equipment called for.

2. The second alternative is, for the various Governments to form what would in effect be a private company or Commission of which the Governments only would be the stock-holders to do the work described in the first alternative and from each subscribing Government the Commission would derive its authority in that country. The Commission being supplied by funds, on some agreed plan of participation by each Government, would be the sole judges of the plant to be constructed and operated, within the scope of the monies put at its disposal, and it would assume the ordinary responsibilities of a board of directors of a public company, carrying out all the necessary functions and periodically reporting results to those who supply the capital. It might be that all plant constructed in any country should belong to that country, and that the capital to be provided by that country should be its proper share depending upon the plant within its own borders.

In addition, the Commission could hire facilities, where economical, from the local administrations in cases where it would not be economical for the Commission to construct its own lines. Such lines could be hired on a permanent or temporary basis. In the first case, they would be paid for at a proper rate per year; in the second case, the Local Authority's lines might be made use of and the compensation to be paid by the through business Commission might be a proper portion of the fee paid by the public. The above is the merest sketch of a scheme, but if it should find acceptance I am ready to put forward for consideration by the proper authorities a plan which I believe will be found to provide a basis on which the countries can be represented on equal terms and by which no unfair burden is placed on any country and I believe such a plan would result in better service and be self supporting.

3. The third alternative is frankly one of a temporizing nature, being intended only to cover a study of this difficult problem; it is that the various operating telephone authorities should form themselves into an association for the purpose of studying this and other matters. Such association might come about gradually if necessary, and regular meetings might be fixed for the purpose of studying a pre-arranged programme, which apart from the larger question as to how the through business should be operated, might include the fixing of standards of measurement, performance and methods to be recommended to all and to be enforceable on those who subscribe to the association.

If I may venture to make a definite suggestion, it is that the telephone authorities of Europe—including the United Kingdom—as telephone operating authorities rather than as Government Departments—should hold an early conference of all the telephone authorities—Companies and Municipalities as well as Government Departments, to study in detail this problem and endeavour to find a solution. I am convinced that unity of control over the through traffic must obtain in the end, but whether the through traffic is handled by one organization or by many, there are matters which urgently require agreement for the improvement of telephony as an efficient agent for service in Europe.

Almost entirely, what has been said is limited to through communication by telephony; this is not because there is nothing to be said regarding local service, but rather because it seemed better to try and focus attention on what at the moment is the greatest telephone problem in Europe, viz., How shall the through business be organized? Fortunately, the solution of this problem has never yet been seriously undertaken, and the whole matter being quite open there are no standing decisions to be reconsidered. The engineering considerations make it plain that the communication which is possible both technically and commercially cannot be established under the present disconnected organization. As with a progressing organism the time has come when the organization must, if it is to remain efficient, change from unicellular to multicellular and the various cells must take up special functions rather than all functions; in that way only can the whole organization make progress.

One way of increasing good-will among nations—especially necessary to be encouraged by all means possible at the present time—is by greater and ever greater intercommunication by all methods. In the telephone we have the most perfect means of communication of which we know, immediate and perfect human speech with all its tones and inflections and the ability by interchange of conversation to remove misunderstandings. If only we will use it, not alone will it benefit the industry of the nation, but we shall be making a definite step towards reducing the international jealousies and fears and increasing the good-will without which there cannot be peace on earth.

Telephone Repeaters

By BANCROFT GHERARDI

Vice President and Chief Engineer of the American Telephone and Telegraph Company

(In the following article the author describes the application of the telephone repeater to the telephone plant. A description of the various types of telephone repeaters which have been employed and their operating characteristics appeared in the August issue of ELECTRICAL COMMUNICATION. —EDITOR.)

PART II—APPLICATION OF REPEATER

FROM an operating standpoint, repeaters are applied to the lines in three different ways. In the first place, there is what is known as the through-line repeater. This is a repeater which is connected into a given telephone circuit and remains there at all times. A second type is the cord-circuit repeater which is a repeater in a cord circuit and may be connected into the line by the operator at any time that she wants it, so that it can be used on a number of different circuits successively. A third type is the emergency repeater, which is made up in portable form, and carried around to meet special and temporary conditions. Figure 11 is a photograph of one view of the emergency repeater. A general idea of the scale of the picture may be obtained by noticing the bulbs in the middle of the top of the panel. The balance of the equipment comprises the associated apparatus—the coils, signaling apparatus, resistances for controlling current, the condensers, and the rest of the apparatus which goes with a repeater installation. It may be noted what a small part the repeater bulb is of the repeater installation as a whole.

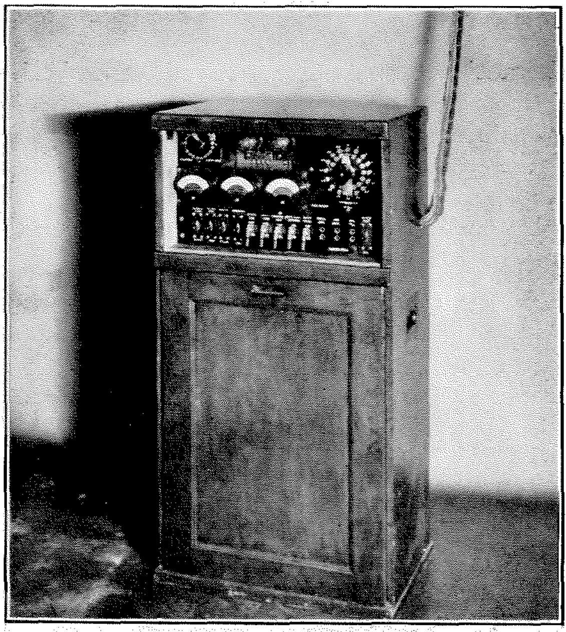


FIGURE 11

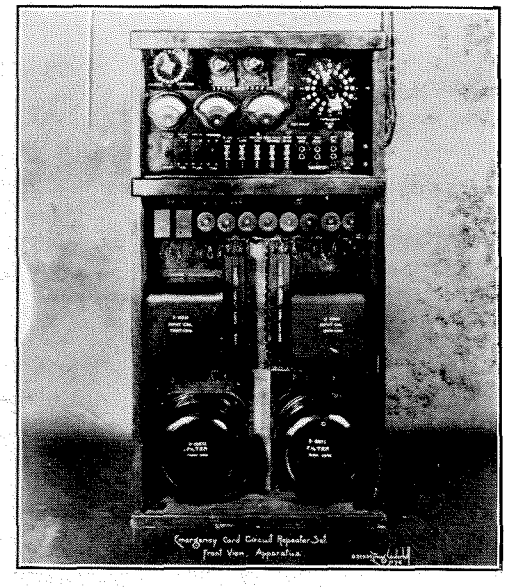


FIGURE 12

Figure 12 is a picture of the front of the emergency set, with the case open showing the apparatus that is used for the purpose of operation and signaling. Figure 13 shows the back of the emergency set.

It will be interesting to note some of the practical applications that have been made of repeaters. There has been a very large number of repeaters installed in the Bell System; there being about two thousand of them working at the present time. On account of the striking character of the Transcontinental Line, and the fact that it was the first line of its kind anywhere in the world, it will be described first.

The Transcontinental Line was built with two terminals at this end, one at Boston and the other at New York, and at the other end it was terminated at San Francisco. The line is operated with thirteen repeater stations, at New York, Harrisburg, Pittsburgh, Beaver Dam, Morrell Park, Davenport, Omaha, North Platte, Denver, Rawlins, Salt Lake City, Winnemucca and Sacramento. With the repeaters in the line has a transmission efficiency of about 4 per cent. Without repeaters the ratio of input energy to output energy would be approximatey $3(10)^{14}$.

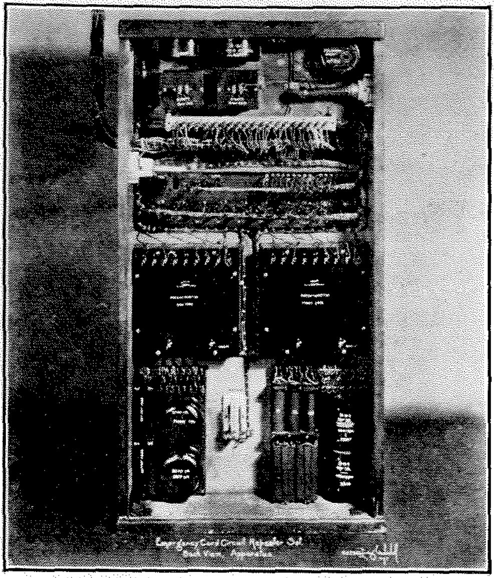


FIGURE 13

It is very interesting at this point to consider what would have been the effect of trying to solve the problem of talking over long telephone circuits by means of loud-speaking transmitters. For a good many years inventors, scientists, promoters—both here and abroad—thought that the solution of the long-distance problem was very powerful transmitters. Such transmitters were designed, and in some cases a good deal of stock was sold based upon them. Now, the first thing we did in studying this repeater problem was to look at it from the broadest possible point of view. Inasmuch as it was desired to give service all over the United States, from one extreme to the other, we asked ourselves the question: "What is the best way to do that?" Consider all the alternatives—loud-speaking transmitters would be one alternative, sensitive receivers would be another, repeaters would be a third, and modified line construction might be a fourth. All of these alternatives were considered, and we decided—rightly, I am sure—that the solution lay in the telephone repeater, so we proceeded to develop the telephone repeater.

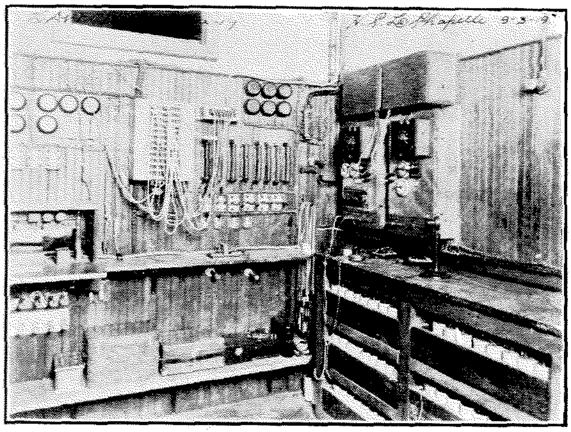


FIGURE 14

To have accomplished on the Transcontinental Line by modified transmitters the results which are obtained by repeaters, it would have been necessary to have a transmitter which would have so great an output of telephone current that it would impress 9,000,000 volts on the telephone line. This, of course, is an absurdly high voltage.

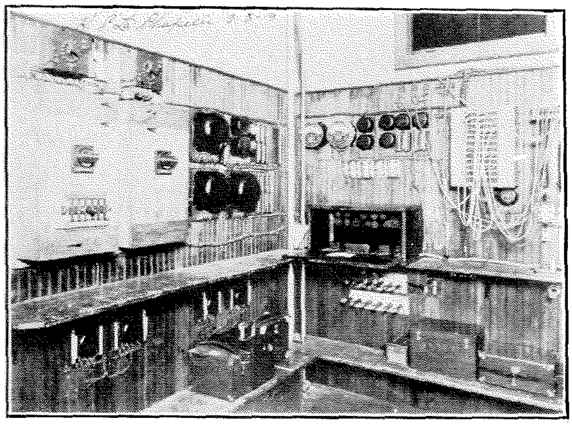


FIGURE 15

Theoretically at least, another way of obtaining that Transcontinental Line conversation would have been, instead of using repeaters, to put more copper into the line. At the present time the Transcontinental Line has per mile about 870 pounds of copper—a little less than 1,000 pounds—so that there are about three and a half million pounds of copper on the line from Boston to San Francisco. To get along without the repeaters, and give the same transmission that is now being given, would require the use of about 36,000,000 pounds of copper in addition to the three and a half million already on the line. Of course, economically, that would be an absurd solution; it would mean that there would be no Transcontinental Line, because no one could afford to pay for transcontinental communication over a circuit weighing about 10,000 pounds per circuit mile.

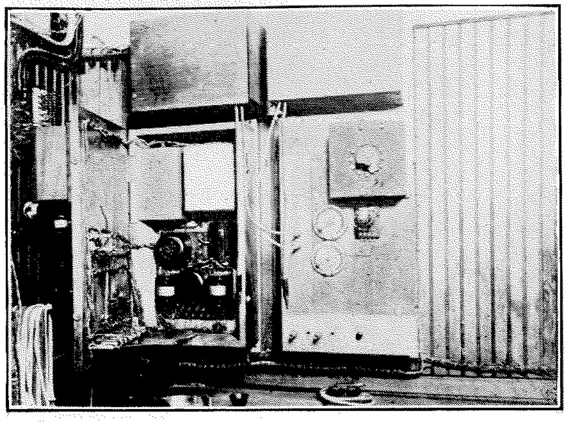


Figure 16

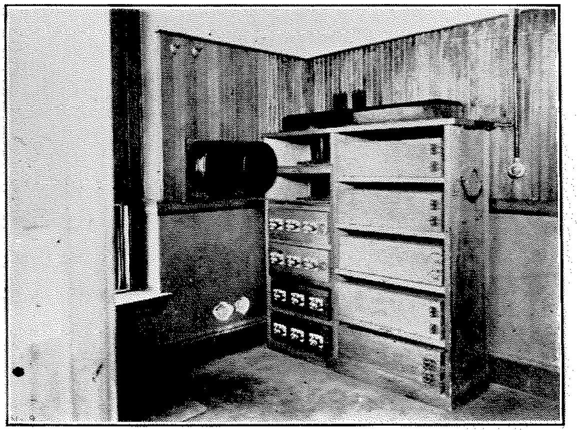


Figure 17

A number of pictures of repeaters on the Transcontinental Line are here illustrated. Figure 14 is interesting. It is the first installation of repeaters made at Denver in connection with the Transcontinental Line. It is interesting chiefly because it is so different from those which are shown in pictures that follow.

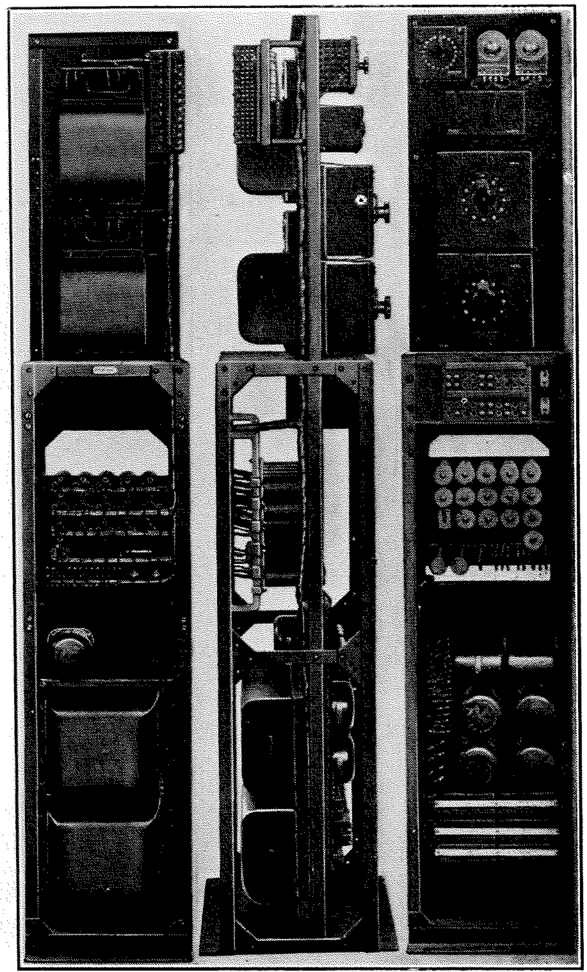


Figure 18

Figure 15 shows some more of that first Denver installation, and Figure 16 is a view closer up of the vacuum tube repeaters in Figure 15. Figure 17 is a view of the artificial lines that balance the regular lines.

Figure 18 is a picture of a modern repeater unit, showing the front view, the back view, and the edge-wise view.

Figure 19 shows an installation of six modern repeaters—not of the latest type, but the type that followed that experimental apparatus.

These are installed in a row with a testing unit in the middle. The testing unit enables us to make the necessary electrical measurements to see that the repeaters are in proper operating condition. Figure 20 is a view of the back of those units showing the mounting of the apparatus at the rear.

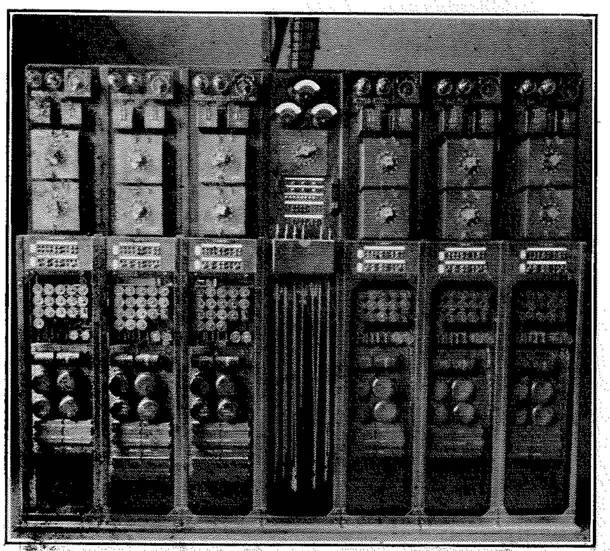


FIGURE 19

Figure 21 is a view of the balancing artificial lines that go with the repeaters shown in Figures 19 and 20.

Figure 22 is a group of vacuum tube repeaters which have been installed in Providence, and which are working on the cable circuits running from Providence. They are very much like the other pictures which have been shown in connection with vacuum tubes. Figure 23 is the rear view.

Figure 24 is a diagram which shows what takes place with reference to the energy on the Transcontinental Line. Of course the amount of energy that is started at the sending end of the line is dependent upon just how loud the person speaks, but, for the purpose of this diagram, we have assumed an average figure, which is 1:100 watt; the energy is attenuated as it passes along the line and amplified at each repeater station, until it finally gets to San Francisco—still up enough to give a thoroughly satisfactory telephone conversation.

The dotted curve on this diagram shows the energy that would have to be in the line if the repeaters were not used, in order to deliver the same energy at the receiving end. This dotted curve is not complete, because at the point marked Boston on the diagram it would have to be 10,000,000 miles high.

Figure 25 shows an installation of mechanical repeaters using the element shown in Figure 3. The elements are at the tops of the panels, and all the rest is the supplemental apparatus necessary in order to make them work. Figure 26 is a back view of the same panels.

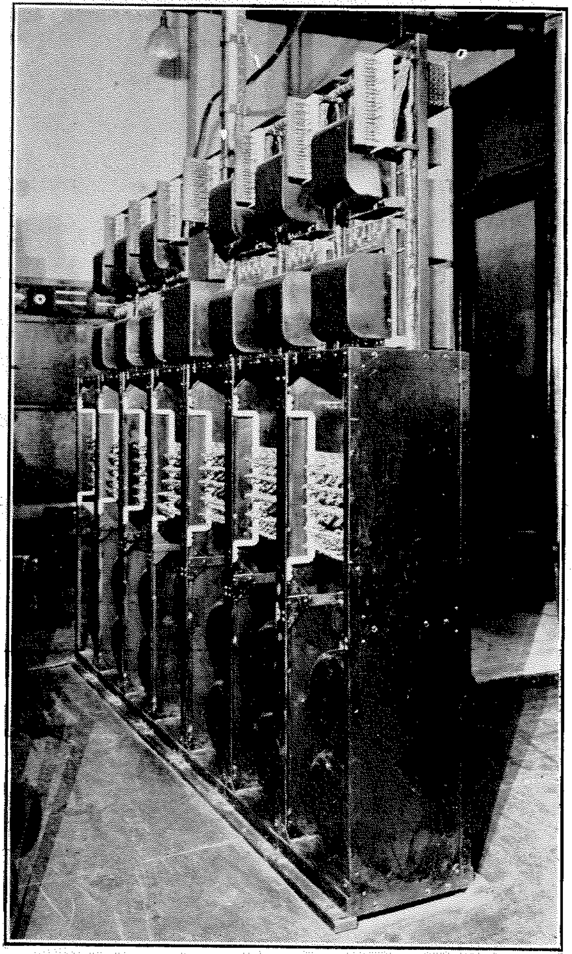


FIGURE 20

Figure 27 is a view of the Princeton repeater station. Princeton is located along the route of the cables extending from Boston to Washington. The first cable was completed in 1914, and since then has been supplemented by other cables. At Princeton, the number of repeaters was so great that it was found worth while to construct a special building merely for their accommodation, and Figure 27 is an external view of that building.

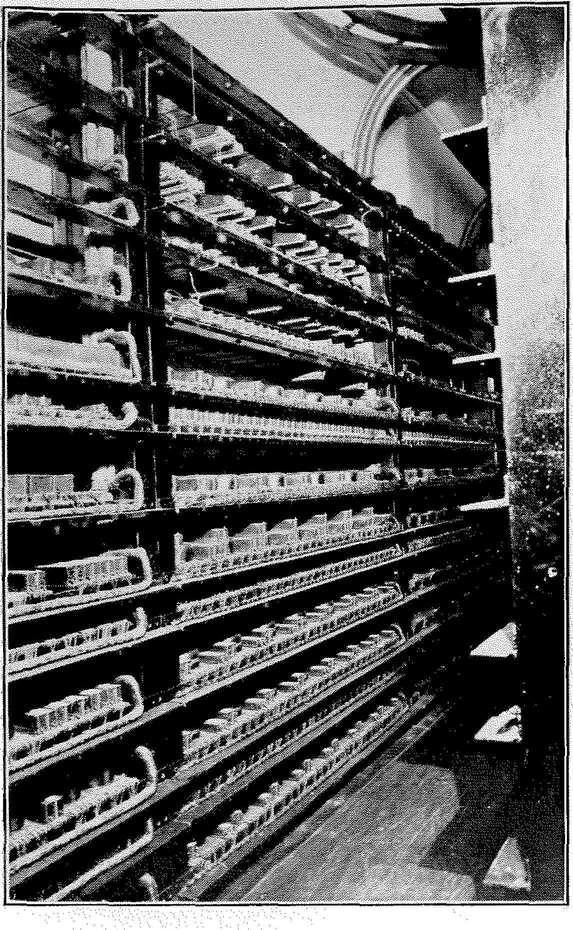


FIGURE 21

Figure 28 is the interior plan of the building, showing the space for the two-wire repeaters, for four-wire repeaters, and room for growth. The test boards, relay racks, fuse panels and distributing frame, are also indicated.

Figure 29 shows one of the two-wire repeater units and Figure 30 one of the four-wire units. It will be noted that there are four repeaters on each panel unit.

Figure 31 shows a bank of two-wire repeater units.

Figure 32 is a view of the cable test board which was shown in the floor plan.

Figure 33 is a view of some supplemental apparatus, showing the fuse panels and the spring jacks by which one can cut in on the various circuits and measure the exact amount of currents which the tubes are taking. That is important in connection with keeping the devices in proper operating adjustment.

Figure 34 is a diagram on which is shown the route of the Boston-Washington cable, and also the route of the first section of the cable which we are now planning to Chicago. That cable is already in place for a short distance out of Chicago, and from Philadelphia to Pittsburgh. By some time next year it will be extended as far as Cleveland.

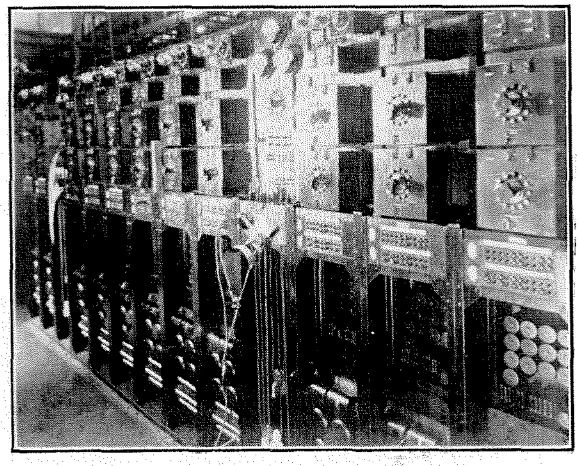


FIGURE 22

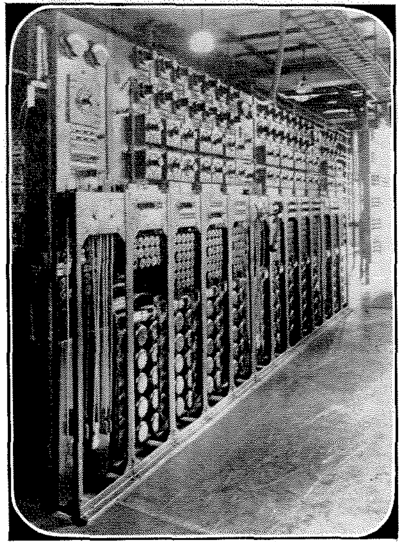


FIGURE 23

Before speaking further about this extension of the cable system to Chicago, I would like to say a little more about the Boston-Washington cable. At the present time, among the most interesting of the circuits that we are working in that cable are the circuits from New York to Washington—2-wire, 19-gauge circuits, 40 pounds per mile. These 19-gauge circuits from New York to Washington present a problem as difficult as the No. 8 circuits from New York to San Francisco. The fact that the distance is so much less is offset by the fact that the wire is so much smaller, and the capacity so much greater—due to being in cable. We have, however, been able to get successful results with those circuits—in fact as good as on the other circuits that were working between New York and Washington. We have cut out about four-fifths, or five-sixths, of the total attenuation of that circuit, so that the actual talk which the public gets over those circuits is as good as though they were only a few miles long.

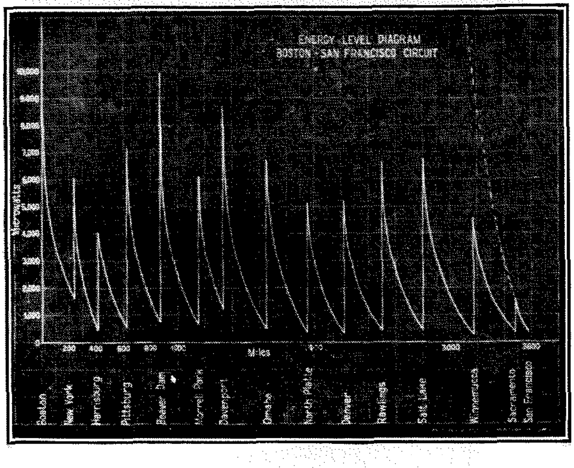


FIGURE 24

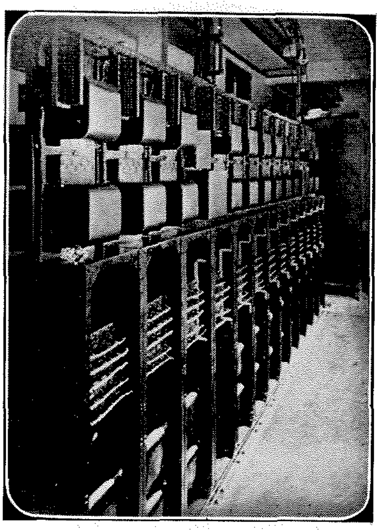


FIGURE 25

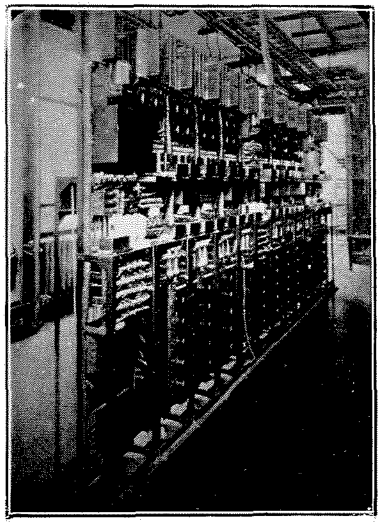


FIGURE 26

Now, in the same way that this cable tied together the most important places in the United States located along the Atlantic seaboard, and provided them with connection which was storm-proof, the cable extending to the west is going to bring all of these Atlantic seaboard points into similar touch with such places as Pittsburgh, Cleveland, Toledo—there is a short spur going up to Detroit—and then with Chicago. Just as soon as we get a second route, which is now under contemplation, from Boston to Albany, Buffalo, Cleveland, Toledo, and so

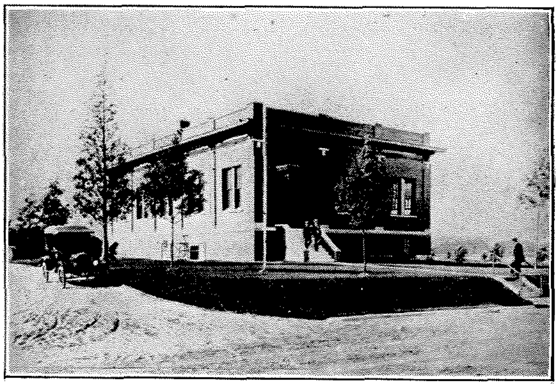


FIGURE 27

on, we will not only have an alternative route from Boston to the West, but we will also have such places as Albany, Syracuse, Rochester, and Buffalo, brought into the combination; so that we will have, by the running of those three cables, and a few spurs—practically all the principal places in a large part of the United States connected together in the cable network.

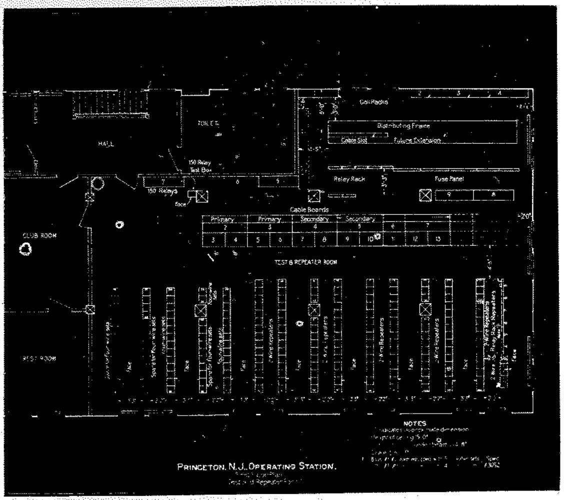


FIGURE 28

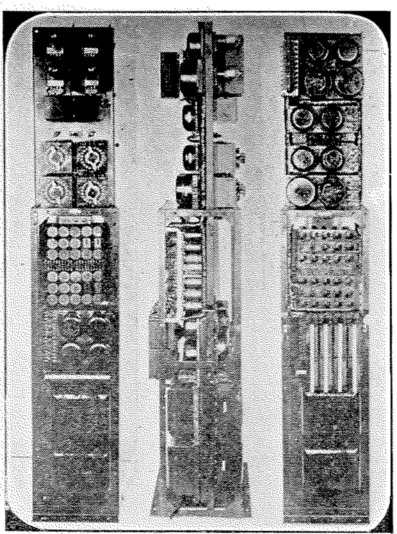


FIGURE 29

The interesting thing is how we are going to get that cable to talk. You will remember it was only a few years ago when there was a terrible dread of even a few miles of cables in a toll line; then it was considered quite an accomplishment, and required quite a good deal of care, to design toll cables to work satisfactorily, from Boston to Lynn, and Salem—and Worcester was a real problem. Then the Chicago-Milwaukee cable came up as a problem; and I can remember very well one evening in Boston

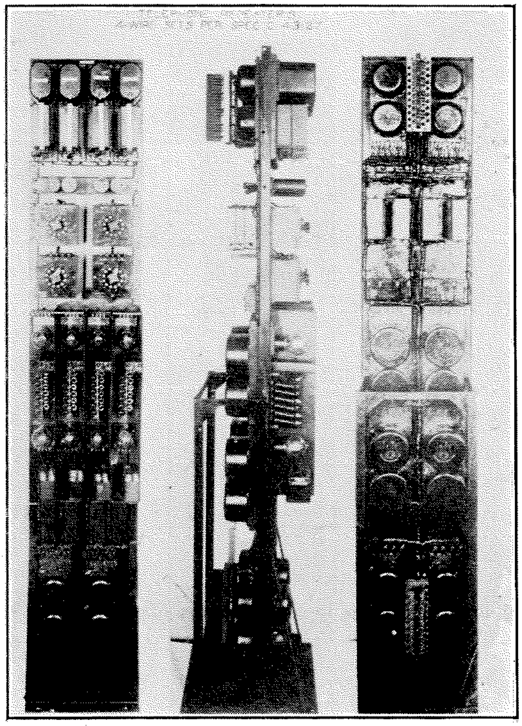


FIGURE 30

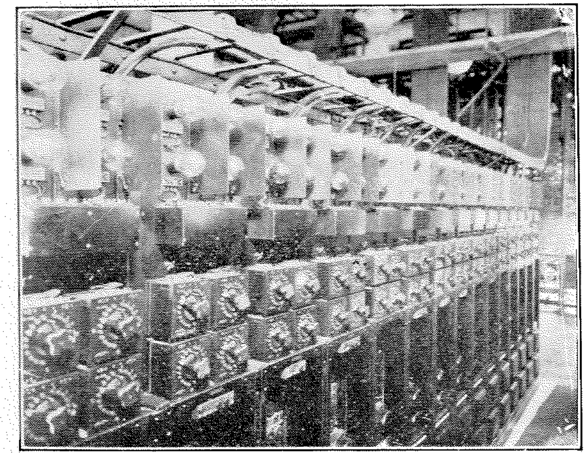


FIGURE 31

when Mr. Hayes, Mr. Fish, then President of the A. T. & T.; Mr. E. K. Hall, Vice-President of the A. T. & T.; Col. Carty, Chief Engineer of the New York Telephone Company, and a lot of other prominent people, were present at a meeting to discuss the cable between New York and Philadelphia—a matter which was then very serious and very difficult. It was just as difficult in those days, as the things which are being done today, starting from the point which was reached in the earlier work.

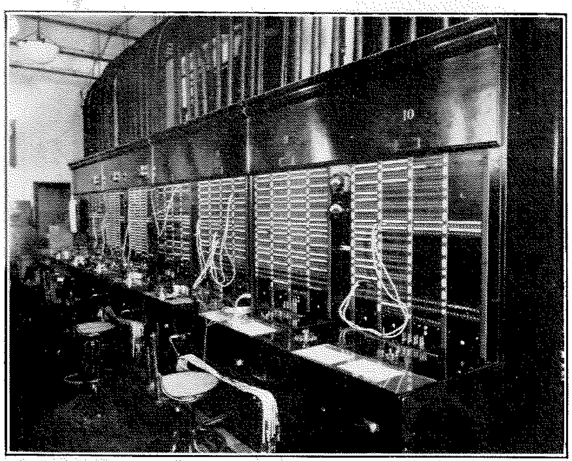


FIGURE 32

One important feature of this project is the use of the four-wire circuit, which has already been described to you. It might seem a waste of material to use two pairs where formerly only one pair was used. It does seem like a waste of material from one point of view, but from another point of view, the original Chicago circuits at 870 pounds per mile circuit, will be replaced by two No. 19-gauge pairs totalling 80 pounds of copper per mile circuit. This is accomplished by applying the four-wire principle.

It is expected that there will be 17 intermediate repeating points between New York and Chicago. The circuit will be so efficient that the energy at the receiving end will be equal to one-tenth of the energy put into the circuit at the transmitting end. It is impossible to tell how much the energy would have to be at the transmitting end if it wasn't for the repeaters; it is in such inconceivably large numbers that it doesn't mean much to any of us. Those that are mathematically inclined may know what 10 to the 44th power is, but it may be illustrated in another way. This circuit will be satisfactorily heard in Chicago if the ordinary telephone transmitter is spoken into in an ordinary tone of voice at Boston or New York. Without the repeaters, if one were to put all of the energy of all of the prime movers in the world into the sending end of that circuit it wouldn't be capable of giving an audible sound at the receiving end. Indeed if one put all of the energy radiated by the sun into the transmitting end of that circuit, he couldn't get anything out of the receiving end. It is almost impossible to conceive of the degree of amplification that is put into that circuit by the 17 repeaters working in series. The amplification in billions can be expressed as a billion, billion, billion, billion million—if that means anything to anybody!—it doesn't mean anything to me.

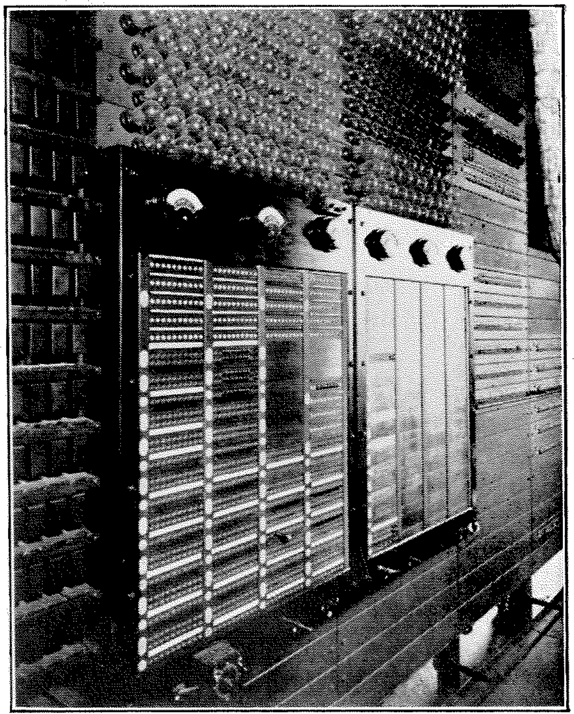


FIGURE 33

But we do know that the circuit will work, because we have set up such a circuit, and talked over it.

The amplification is so great as to give a number of very interesting results—one of which may be mentioned. In the circuit there is a very nice balance between an enormous loss and an extremely great amplification; that is, the amplification is just a little bit less than the loss. Those two things are so very large and the difference is relatively so small, that there is a big change in the efficiency of the circuit with ordinary changes in temperature, such as are brought about by the sun shining on the cable for a few minutes. It is necessary therefore to have arrangements in the circuit for automatically regulating the amplification in the repeaters, to compensate for the changes in energy loss due to the changes in resistance of the copper wires just from heating them a little bit or from cooling them off a little bit. With a change of temperature which can easily take place in passing from a general cloudy to a sunny condition, the amount of energy received at the other end without compensation would be changed by as much as a thousand times. If the change were upwards, without compensation the circuit would talk so loudly that it would sing, and, if the change were downwards, one couldn't hear on the circuit at all.

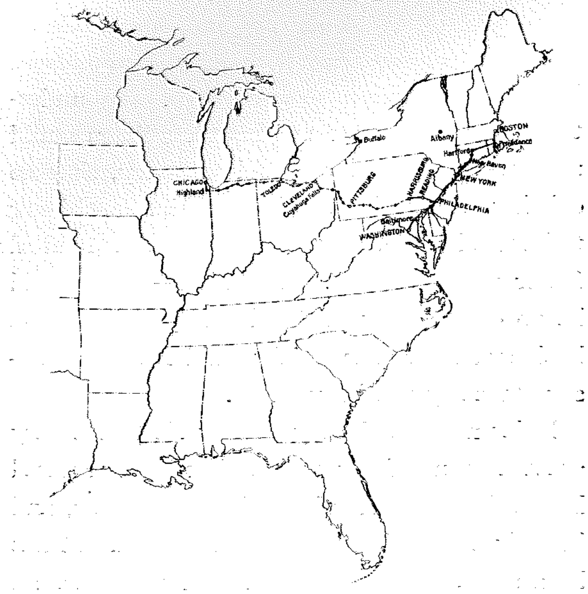


FIGURE 34

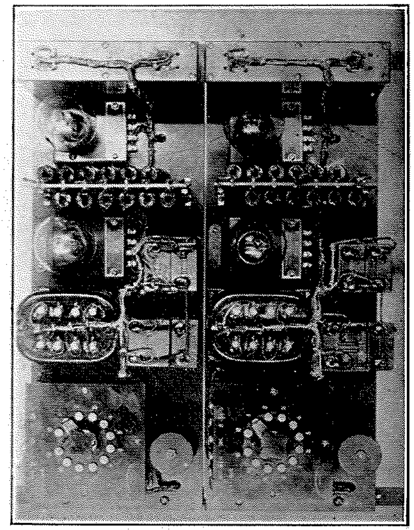


FIGURE 35

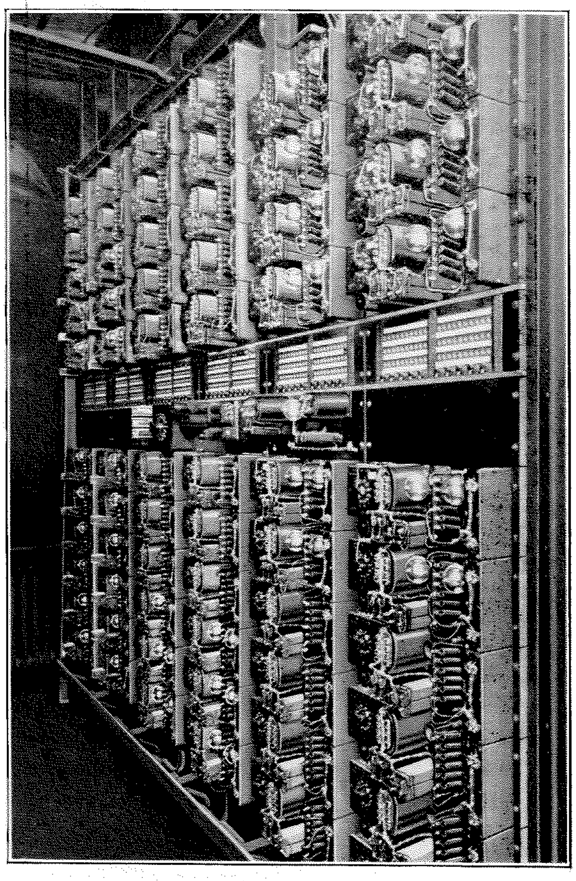


FIGURE 36

You will be interested in some pictures of the type of repeater by which these marvelous results are accomplished. Figure 35 shows a front view of two of these repeaters. Figure 36 shows the front of a bank of sixty repeaters as actually installed on this cable. The form of assembly is such as to lead to large economies in space and operating costs. Figure 37 is a back view of a bank of the same type of repeaters.

Another very interesting development in the extension of telephone circuits which has been made possible by means of telephone repeaters is the cables which have been laid between Key West and Havana, a distance of about 125 miles. We have laid in conjunction with the Cuban Telephone Company, three one-conductor cables for this service. Each of these cables is operated as a telephone circuit with superimposed telegraph facilities. The mechanical conditions were largely controlling in the design of those cables. In the first place, we were not able to deal with the repeater problem in our usual way—for there is a stretch of 125 miles with no chance to put repeaters in at intermediate points, as we should like to. In the next place, the water is over 6,000 feet deep at the deepest point and a considerable part of it is over a mile deep. It is so deep that the mechanical conditions are controlling, and we cannot safely use an ordinary type loading coil, or even special types of loading coil. However, repeaters located at the cable terminals have been designed so that we are able to get very satisfactory talking circuits across on those cables. With the amplifiers that we use, the energy at the receiving end of the cable is about one-sixth of the energy put into the cable at the transmitting end; without the amplifiers, it would be one-seven-hundredth. To undertake to use those cables without amplifiers would require a cable so large, and so heavy, that no existing cable-ship could have the machinery to handle it; also, there would be the economical factor to take into consideration—there would be no opportunity to develop any considerable business at the rates which would then be necessary to carry the investment.

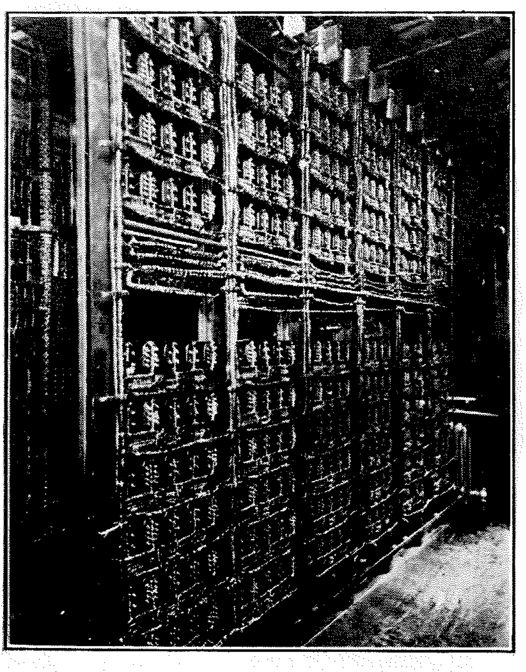


FIGURE 37

Reference has already been made to the very great amount of work which has been required in the development of telephone repeaters, and in their application to the Bell System. It has been necessary to develop not only the vacuum tubes and coils which are used to reproduce the speech currents faithfully, but also the many other pieces of apparatus and circuits which are essential if the repeater is to operate satisfactorily in the existing telephone plant. I refer to signalling equipment—the apparatus necessary to enable one operator to signal another operator even though the circuit connecting them contains several telephone repeaters. We must also provide means for permitting the telegraph currents to pass through the repeater stations uninterrupted. Like all machines, repeaters get out of order occasionally. So as to be able to diagnose and remedy the trouble promptly, we have equipped them with apparatus so that the different parts of the circuit can be measured. Apparatus for determining the amount the voice is magnified is also used so that we are able to determine whether the subscriber's voice when received at the distant station will be loud enough to be easily understood. It will be apparent that even after we had perfected our vacuum tubes and coils, it was still necessary for us to design numerous other devices to enable the repeater set to operate to the greatest advantage.

Notes on Loaded Long-Distance Telephone Cables in Europe

By P. E. ERIKSON

European Assistant Chief Engineer, International Western Electric Company

IN the following notes an attempt has been made to outline, briefly, some of the achievements in the field of long-distance cable telephony in Europe. The extraordinary development in the telephone art witnessed in the United States during the past decade, has had a healthy reaction in the more progressive European countries.

But for the intervention of the War which, of course, severely retarded all activity in commercial telephony, progress in Europe today would undoubtedly have been much more advanced. However, since the Armistice and more particularly since the early part of 1919, the demand for reconstruction of, and added facilities in the long-distance lines has been greater than ever. This sudden demand for a large number of telephone circuits could not be carried out efficiently by running more wires on existing poles, or by erecting additional pole lines where the existing ones were filled to their ultimate capacity. Nor would it have been good engineering practice to adopt such a course, when it was known that the unreliable open wire lines could profitably be replaced by the modern telephone cable, which is practically immune to the vagaries of any climate.

It is not within the scope of these notes to describe the development of the paper-insulated loaded telephone cable which, by virtue of numerous refinements in design and manufacture, permits of the most efficient use of the conductors contained in the relatively small space afforded by the surrounding lead sheath.

By way of illustration it may suffice to say that, in a lead pipe having an internal diameter no greater than 6 cm., this cable can carry 300 pairs of telephone wires yielding 450 telephone circuits which, by the aid of amplification through loading coils and thermionic repeaters, are rendered far more efficient than the old style open wire circuits.

It is not surprising, therefore, that when this demand for additional circuits arose, several of the progressive Telephone Administrations in Europe turned their attention to this means of meeting it. The countries which first gave this question serious consideration were Great Britain, Holland, Italy, Sweden and Switzerland. Other countries have, to a lesser extent, followed the lead of those mentioned.

GREAT BRITAIN

In a country as densely populated as Great Britain and having its more important commercial and industrial centers scattered over a relatively small area, the need for sufficient long-distance telephone lines is easily understood.

As a matter of fact, Great Britain has been the pioneer in long-distance cable development in Europe. Loaded long-distance cables were installed for commercial use as far back as 1908 when the first Liverpool-Manchester cable (62km.) was installed. This cable was not designed or balanced for loaded phantom operation, but the pairs were loaded with toroidal iron core loading coils of Western Electric design. Previously the British Post Office had experimentally loaded some circuits with solenoidal, air-core coils, but owing to difficulties with insulation and external disturbances these coils were abandoned in favor of the toroidal iron core type.

The introduction of phantom loading in long-distance cables in Great Britain took place in 1913 when the laying of the London-Birmingham cable commenced. This cable, which is 176 km. long, contains 52 physical and 26 phantom circuits, a large proportion of the latter being loaded at the outset and provision made for the insertion of further phantom loading coils later. When this cable was designed the use of telephone repeaters was not contemplated; the dimensions of its copper conductors, therefore, in order to give the desired transmission, had to be relatively large, the largest being 3.48 mm. in diameter. The transmission equivalent, between test-boards, of this type of circuit (loaded) is 4.2 miles of standard cable ($\beta l = 0.46$). At the time of its completion the London-Birmingham cable had no equal in Europe in magnitude or efficiency.

The cable was extended to Liverpool some time after the completion of the London-Birmingham section and in 1916 Great Britain had the longest phantom-loaded toll cable in Europe (320 km.) in successful operation.

About the time of the installation of the London-Birmingham-Liverpool cables, several other long-distance cables were being laid. Among these may be mentioned the Leeds-Hull (95 km.), the London-Brighton (85 km.), London-Colchester and the Birmingham-Leeds (173 km.) cables.

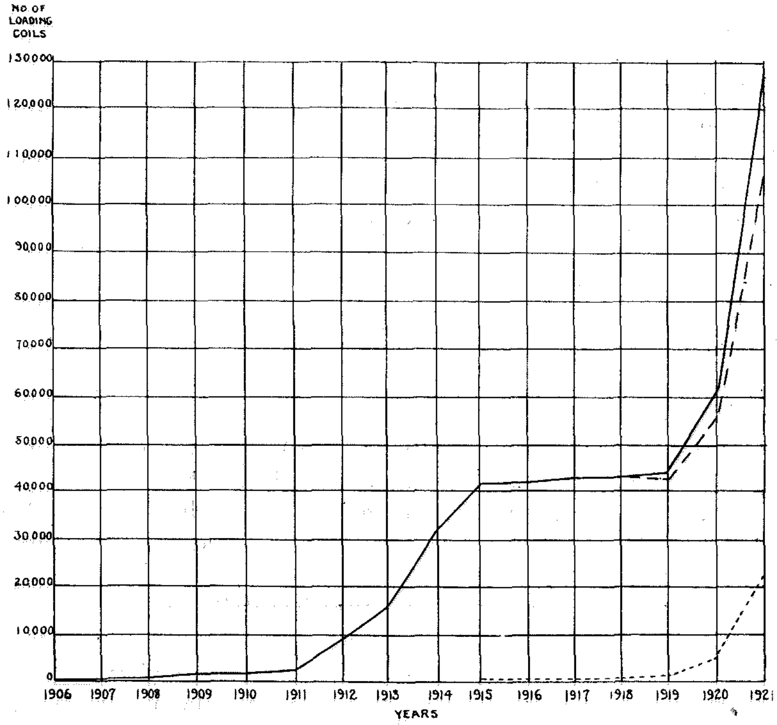


FIGURE 1

NUMBER OF WESTERN ELECTRIC LOADING COILS INSTALLED IN EUROPE 1906-1921

——— Total installed. — — — United Kingdom. - - - - - Continental

These cables formed but a small part of an elaborate long-distance cable programme which, while planned before the War, could not be carried out during that time. Thus it was contemplated to link up several of the commercial centers in the Midlands with London. London-Manchester and London-Leeds cables were to be two of the main arteries from which extensions were planned in several directions.

When the War ended, the British Post Office actively turned their attention to carrying out the aforementioned programme, which was augmented to include several other main routes, notably among others the London-Bristol (198 km.), London-Southampton (137 km.), London-Manchester (300 km.), London-Leeds (324 km.).

A complete list of all the long-distance cables in Great Britain would occupy more space than these notes call for. Some idea of the amount of loaded long-distance cable laid to date may be formed from the fact that 257,600 circuit-kilometers of loaded cable are now in service, which is considerably more than one-half the total of all loaded toll cable in Europe (excluding Germany, from which statistics are not available). Approximately 160,000 additional circuit-kilometers of loaded cable are in course of construction.

In the commercial application of loading coils to telephone cable circuits, Great Britain has been the leading country in Europe. The British telephone authorities who, up to the end of 1911 consisted of the Post Office and the National Telephone Company, were fully alive to the economies which could be effected by the use of loading coils.

The long-distance lines in Great Britain have been under the control of the Post Office since 1896, while the National Telephone Company's activities were confined to the local areas in London and to the provinces.

The National Telephone Company began using loading coils as early as 1906, when some of their important cable junction circuits were loaded.

The growth of the commercial development of loading coils may be seen by reference to the Curve Chart (Fig. 1) which shows the number of loading coils manufactured by the Western Electric Company and furnished to the Telephone Administrations in Europe during the years 1906–1921. It will be noted that by far the greater proportion of the coils were used by the British Telephone Authorities.

HOLLAND

A study of the telephone statistics of Holland reveals some interesting facts. During the ten-year period 1904–1914, the number of telephones per 100 inhabitants increased threefold. In 1914 there was one telephone per 70 inhabitants, while by December 30, 1920, there was one for every 42 inhabitants. Furthermore, the figures show that about one-half of all the telephones in the country are to be found in the three largest cities, Amsterdam, Rotterdam and The Hague, in each of which the local service is operated by the municipality.

In 1911 it became quite evident that the telephone traffic between the cities was growing at such a rate that the aerial lines could no longer provide the desired circuit facilities.

The Dutch Telegraph Administration which operates the long lines therefore decided to lay a cable between these cities, the scheduled date for the commencement of the work being set for 1914. The ducts which were to hold the cable and the manholes for jointing the cable and providing accommodation for loading coils had been constructed and everything was ready for the cable installation. The outbreak of the War, however, stopped all activities in this line and it was not until 1919 that the laying of the cable began. Once the laying of the cable had commenced work proceeded apace and commercial operation was begun in September, 1920. The cable is 90 km. in length and contains 180 loaded telephone circuits, of which 60 are loaded phantom circuits.

In the Amsterdam-Rotterdam cable a total of 180 loading coils had to be installed at each point and for convenience, these were mounted in four containing cases per point. The cable itself is laid in one channel of a 4-way concrete duct, which extends practically the whole distance.

This cable, while serving the three principal cities, also provide circuits for intermediate towns. Some high grade circuits (21 in number) are included in the cable for connection to other circuits in the toll line system of the country. These circuits consist of 2 mm. conductors, the side circuits of which give an over-all transmission equivalent for 90 km. of 5.7 standard miles ($\beta l = 0.62$), while the phantom circuits equate 4.6 standard miles ($\beta l = 0.5$) for the same distance. The cable also contains 1.3 mm. and 0.9 mm. conductor circuits which are intended for terminating and intermediate traffic.

The next cable to be installed was the Amsterdam-Utrecht cable, which may in a sense be considered as an extension of the Amsterdam-Rotterdam cable. Eventually, this will be extended to Rotterdam (over Utrecht), thus providing direct traffic channels between Amsterdam and Rotterdam. This cable, which is of the armoured type and placed directly in the ground, is about 42 km. in length and contains 128 loaded telephone circuits, of which 39 are loaded phantom circuits. This work was completed and some of the circuits put into service in May of this year.

In addition to sixteen quads of 2 mm. conductors, affording 48 telephone circuits of the same efficiency as those referred to previously on the Amsterdam-Rotterdam cable, there are 23 quads and 11 pairs of 1.42 mm. loaded conductors, yielding 80 circuits. The phantom circuits of this grade have a transmission equivalent of 3.8 miles of standard cable ($\beta l = 0.41$) between test boards, while the side circuits equate 4.8 miles of standard cable ($\beta l = 0.52$) for the same distance.

A third important link in the long-distance cable system in Holland is the Rotterdam-Dordrecht cable. This cable is identical, so far as construction and circuit facilities are concerned, with the Amsterdam-Utrecht cable. The length is about 20 km. All the cables referred to were manufactured by the Western Electric Company.

Owing to the large number of river crossings and waterways which had to be traversed, it was necessary to install several sections of submarine cable in all of the long-distance cable lines. The construction of these submarine sections does not differ greatly from the land cables, except that a somewhat heavier iron wire armour was applied over the lead sheath as a means of protection. The conductors were insulated with paper, just as in the land cables, and the electrical constants maintained uniform throughout.

The loading plan adopted for all of these cables called for spacing of the loading coils at distances of about 3 km.

In the case of the Amsterdam-Rotterdam cable the loading coils are buried in the ground a short distance from the jointing chambers where the ducts are terminated. They rest on a bed of concrete and are enclosed by wooden sides. The loading coil stub cables are armoured by steel flexible tubing, at the end of which a straight joint is made to short lengths of armoured auxiliary cable connecting with the main cable in the jointing chambers. On later cables in Holland large concrete manholes were constructed to accommodate the pots and connecting cables. The total number of circuit-kilometers contained in these three loaded cables amounts to 25,385.

During the present year the system is to be extended. Cables of similar construction to the Amsterdan-Utrecht cable are to be installed between Dodrecht and Breda, between Breda and Roosendaal and between Breda and Tilburg. The extension of the cables to the places mentioned is in line with the aim of the Dutch Telegraph Administration's desire to provide good circuits for connections to Great Britain through channel cables. It is also proposed to install another cable between Rotterdam and Amsterdam, as the present cable does not supply sufficient circuits for future requirements.

ITALY

In March, 1913, the Italian Parliament passed a law which will have a far-reaching influence upon the long-distance telephone development in that country. The Ministry of Posts and Telegraphs had prepared an estimate covering the supply and installation of a loaded cable system which is intended to link up the more important industrial centers in Northern Italy, Lombardy, Piemonte and Liguria. The estimate covering this project called for an expenditure of 50 millions of lire (two millions of pounds sterling, or nearly ten millions of dollars, at par). Since Milan, Turin and Genoa, the three largest cities in the districts mentioned, would receive the greatest benefits of this cable system, the project became generally known as the "Milan-Turin-Genoa cable." The combined length of this cable system was to be about 308 km., representing approximately 36,200 circuit-kilometers of loaded cable. Before proceeding with the project it was necessary to obtain parliamentary sanction, which, according to established practice in Italy, takes the form of a law. As already mentioned, the law authorizing the Ministry of Posts and Telegraphs to undertake the work, was passed in March 1913.

Certain preliminary work had to be done by the Ministry and other Governmental Departments consulted, before tenders for the execution of the work could be issued in final form. The work was therefore delayed and then the outbreak of the War suspended further activities.

In 1919 the project was revived and reexamined in the light of the most modern developments in long-distance telephony. Prior to 1914 the use of loaded cables with fairly heavy copper conductors had been contemplated, but with the knowledge of the successful introduction of the thermionic telephone repeaters on loaded cables, permitting of a reduction in the size of the copper conductor required, the Technical Director of the Telegraph Administration deemed it advisable to revise the project of 1913.

Although the plans for the revised project were prepared in 1919, a new factor entered into the problem; safeguarding the cable against power interference.

The hydroelectric development in Northern Italy has, in recent years, reached a stage unequalled in most European countries. High tension power lines, supplying electricity for industrial purposes and for electric railroads, are erected both along highways and across country. It was therefore necessary to select a route which would insure immunity to disturbances from these lines. An investigation was carried out during 1920 for this purpose and alternative routes selected.

In 1921 the contract for supplying the cables, loading coils, telephone repeaters and associated apparatus was awarded the Societa Italiana Reti Telephoniche Interurbane, which is affiliated with the International Western Electric Company.

The cable will be laid underground in ducts practically the whole distance and the work is now about to start.

SWEDEN

In point of telephone development Sweden ranks second among the European countries. There is one telephone per 14 inhabitants and this is surpassed only by Denmark, where the ratio is 1 to 12. Geographically, Sweden covers so great an area that the long-distance line development in recent years has assumed considerable proportions.

During the War the traffic on the main toll lines increased to such an extent that the Swedish Telegraph Administration decided to investigate ways and means of providing adequate circuit facilities to care for present and future demands. The traffic study made showed clearly that the best solution of the problem was to put down cables containing sufficient circuit facilities to provide for growth up to 1931.

The two largest cities in the Kingdom, Stockholm, the capital on the east coast, and Gothenburg, the most important commercial city on the west coast, were in greatest need of improved communication. The Administration therefore decided that this route should be the first one cared for. Tenders were asked and in August 1919 the International Western Electric Company submitted a cost estimate for the supply of a modern telephone cable, equipped with loading coils and telephone repeaters. The estimate was approved by the Telegraph Administration who, through the Treasury, asked Parliament for the necessary funds. On January 16, 1920, the Parliamentary Proposition was introduced and shortly afterwards passed and approved. This cable, while primarily intended to serve the two cities located at its termini, links up several important towns on the way. From these centers the traffic will be switched to open wire lines for other parts of the country.

The total length of the cable is 540 km. The number of conductors varies on different sections, depending upon the circuit demands between intermediate points. The aggregate length of loaded circuit in the whole cable amounts to about 96,650 circuit-kilometers. There are eight repeater stations (including those at the terminals) with an initial equipment of 298 thermionic telephone repeaters. The repeater stations are designed to permit of future extensions to more than double their present capacity. All of the cable, loading coil and repeater equipment has been delivered and the constructional and installation work is nearing completion. When completed the Stockholm-Goteborg cable will be the longest phantom loaded and repeatered underground cable in Europe.

SWITZERLAND

Although relatively small in area Switzerland, owing to its central location, occupies an important position with reference to long-distance telephone communication in Europe. In point of telephone development the country ranks fourth in Europe, the latest figures available showing that there is one telephone to 23 inhabitants. In every department of telephone service Switzerland has always been to the fore, utilizing the latest improvements in the art. Several international telephone lines traverse the country and the Swiss Telegraph Administration has always coöperated with the neighboring countries by giving over their best circuits for this purpose.

Experience had shown that, owing to the rapid development of electric railway traction and supply of electric power for industrial purposes, difficulties would arise as a result of unavoidable disturbances from these sources. In order to overcome this trouble the Administration decided to go in for underground cable in their long-distance lines.

Broadly speaking, there are two main routes over which the toll traffic passes: north-south and east-west. The former originates at Basel, runs through Olten, Lucerne, past the Lake of the Four Cantons through the Gotthard Tunnel, and over Lugano (Italian part of Switzerland) to Northern Italy. This affords communication from the Northern Countries (England, Belgium, Holland and Germany) to Italy. The west-east line begins at Geneva and runs through the middle part of Switzerland to Lake Constance, linking up several important cities, such as Lausanne, Berne, Zurich and St. Gall.

Realizing the importance of these traffic routes the Swiss Telegraph Administration started their long-distance cable system by replacing the overhead lines along these routes. One of the more important cables laid is the Lucerne-Attinghausen cable, which passes through Arth-Goldau at the foot of the Righi, along the Lake of the Four Cantons in the direction of the Gotthard Tunnel. Although the present cable is relatively short (54 km.) it forms the backbone of the north-south route, which will eventually be placed in cable as far as practicable.

The number of loaded conductors contained in this cable varies along the route, the maximum number of circuits in any portion being 93, including the phantoms which are also loaded. The cable is of a composite type, containing 1.5 mm. and 1 mm. conductors, according to the service need. The over-all transmission for the 54 km. length is about 8.5 standard miles ($\beta l = 0.93$) for the 1 mm. side circuits and 7 standard miles ($\beta l = 0.77$) for the phantom circuits. The transmission equivalents for the 1.5 mm. circuits are about 45% lower than the corresponding 1.0 mm. circuits.

In order to provide better service between Zurich and Southern Switzerland a cable of some importance is now being laid from this town to Arth-Goldau over Zug, a distance of 42 km. Constructionally this is identical with the Lucerne-Attinghausen cable, the kilometric transmission efficiency being the same.

Another cable which will be of considerable importance is the one which is about to be laid in the Simplon Tunnel. The telephone traffic from Western Switzerland to Italy has been carried over the lines erected along the Rhone Valley and, until quite recently, the circuits extending to Italy were accommodated in the tunnel over a 7-pair Krarup cable which was laid in the railway tunnel and opened for traffic in 1905.

The telephone service requirements over this route have increased to such an extent that the Swiss and Italian Telegraph Administrations decided to install a new cable of a more modern type. The actual length of this cable is 22 km. There are 20 pairs of 1.0 mm. conductors, arranged for phantom working, side and phantom circuits all being loaded, making a total of 30 circuits.

Geographically, one-half of the Simplon cable is to be laid on Swiss territory and for national reasons had to be made in Switzerland. The cable works at Cortaillod, affiliated with the International Western Electric Company, were awarded the contract for the manufacture of this cable which is now under way. This Company also made the Lucerne-Attinghausen and the Zurich-Arth cables previously described, the loading coil equipment being supplied by the International Western Electric Company. The Italian portion of the Simplon cable was manufactured by the Societa Italiana Reti Telephoniche Interurbane.

An interesting construction feature may be mentioned. Instead of using multi-way ducts, the Swiss Administration employs a single-way duct on all the larger routes. This duct is 25 cm. in diameter and accommodates several lightly armoured cables, which are drawn in as required.

In conclusion, it should be mentioned that in planning their long -distance cable circuits, the Swiss Administration has followed the most up-to-date practice of using small gauge conductors, loaded and intended for operation with telephone repeaters. The programme is being carried out systematically and as rapidly as economic conditions warrant.

The Physical Characteristics of Audition and Dynamical Analysis of the External Ear

By R. L. WEGEL

Engineering Department, Western Electric Company

1. *Introduction.* It has become important in the design and development of telephone apparatus and circuits to know quantitatively the various functional characteristics of the ear since the ear is an important dynamical unit in the long series of vibration transmitting apparatus constituting a telephone system. A complete analysis of this problem involves not only the properties of the physical circuit, but also the characteristics of the ear and voice and of the air passages between the mouth and transmitter and between the ear and receiver. It is the purpose of the present paper to discuss some of the characteristics of the ear and its outer air passages.

Much has been learned about the normal ear by the investigation of the characteristics of abnormal ears. This has incidentally had an application to otological diagnosis and the design and building of amplifying apparatus for the deaf.

This paper is a summary of the conclusions reached to date regarding the absolute sensitivity of normal and abnormal ears, the maximum sound to which the ear can accommodate itself, the much discussed points of "upper and lower frequency limits of audition," the "quality" of audition, a brief mention of the binaural sense and the principles of rigorous dynamical analysis of the ear as a mechanism. A brief description of the apparatus used is also given.

The function of the auditory sense is to detect sound of various kinds and wave shapes varying over a range of pressure on the ear drum of from about .001 to 1,000 dynes per cm.^{2} and over a considerable part of this range to differentiate with certainty between complex sounds so nearly alike that no existing physical apparatus can separate them. The binaural feature adds a sense of orientation with respect to a source and uniform sensitivity for sounds approaching from different directions. The abnormal auditory sense may be regarded as lacking more or less in (a) range of sensation (frequency and intensity); (b) quality of sensation in various regions of the range; (c) the binaural sense. Apparatus and methods have been developed by means of which the outstanding features of these functions can be measured and to a limited extent compensated for.

2. *Minimum Audibility.* Fig. 1 shows a plot of the logarithmic average of minimum audible pressure on 72 normal ears taken throughout a range of frequency from 60 to 4,000 cycles.[1] Both the intensity and frequency scales are logarithmic. Although all skew errors in the determination of the average curve have not been eliminated, an investigation has shown that they are so small as not to affect the utility of the curve for the purpose of measuring deafness. Among the errors which obviously tend to raise this curve might be mentioned, noise in the observing room, abnormality of hearing, lack of attention, and low mentality of the observer. Care was taken to reduce these errors to a minimum without actually making separate quantitative measurements of each of them on a rigorous statistical basis.

The statistical deviation from the mean varies irregularly with frequency; very likely this is due mostly to the external anatomical variations in ears which cause deviations in the dynamical constants of the transmission system from sound source to the ear drum. The dotted lines following the curve of minimum audibility represent approximately the "standard deviation."

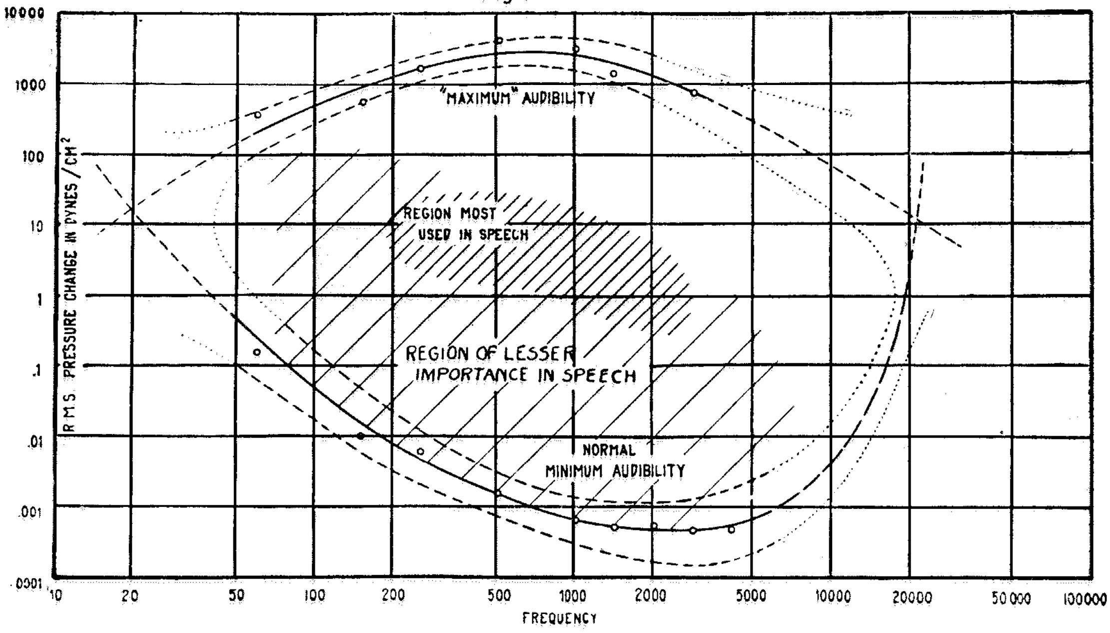
3. *Maximum Audibility.* The curve marked "Maximum audibility" represents the logarithmic average pressure on 48 normal ears required to produce the sensation of feeling. This represents the threshold of feeling in the same way that the minimum audibility curve repre-

---

[1] This curve has already been published: The Frequency Sensitivity of Normal Ears, by H. Fletcher and R. L. Wegel, *Proceedings of the National Academy of Sciences*, January, 1922, and *Physical Review*, June, 1922.

sents the threshold of audition. A sound much louder than this is painful. The measurements were taken through a range of from 60 to 3,000 cycles. The standard deviation lines are also given from which it will be seen that this curve is quite as definite as that of minimum audibility. While this point of feeling probably has no relation to the auditory sense it does serve as a practical limit to the range of auditory sensation. A few observations indicate that people with abnormal ears have a point of feeling sound which is not greatly different from that of normal ears, but this, of course, depends on the type of abnormality. The intensity for feeling is about equal to that required to excite the tactile nerves in the finger tips.

Fig. 1



AUDITORY SENSATION AREAS

4. *Lower and Upper Frequency Limits of Hearing.* The curves of minimum and maximum audibility in Fig. 1 will be seen to have been extrapolated to the points of intersection at high and low frequencies. The feeling sensation in the middle range of frequency is first a tickling sensation and then becomes acutely painful as the loudness is increased. As the frequency is decreased the sensation of feeling becomes milder until frequencies around 60 cycles it is sensible as a flutter, but still quite different from the sense of audition. As the frequency is still further decreased to a point where the hearing and feeling lines appear to intersect, it is difficult to distinguish between the sense of hearing and that of feeling. The low point of intersection of the two normal curves of minimum audibility and feeling sense may, therefore, be taken arbitrarily as the lower tone limit of audibility. For frequencies lower than this it is easier to feel than to hear the air vibration. The point of intersection cannot be determined by direct observation due to the difficulty in distinguishing between the two sensations. A similar intersection of the two curves occurs at some very high frequency. Sound waves of frequencies below the lower intersection and above the upper intersection are more easily sensed by feeling. Sound waves between these limits are more easily sensed by audition.[2]

This suggests a rational way of defining and determining the two frequency limits of audibility. Measurements of these limits which have been made in the past are questionable because the intensity factor has been neglected. At the lower limit of audibility the excursions

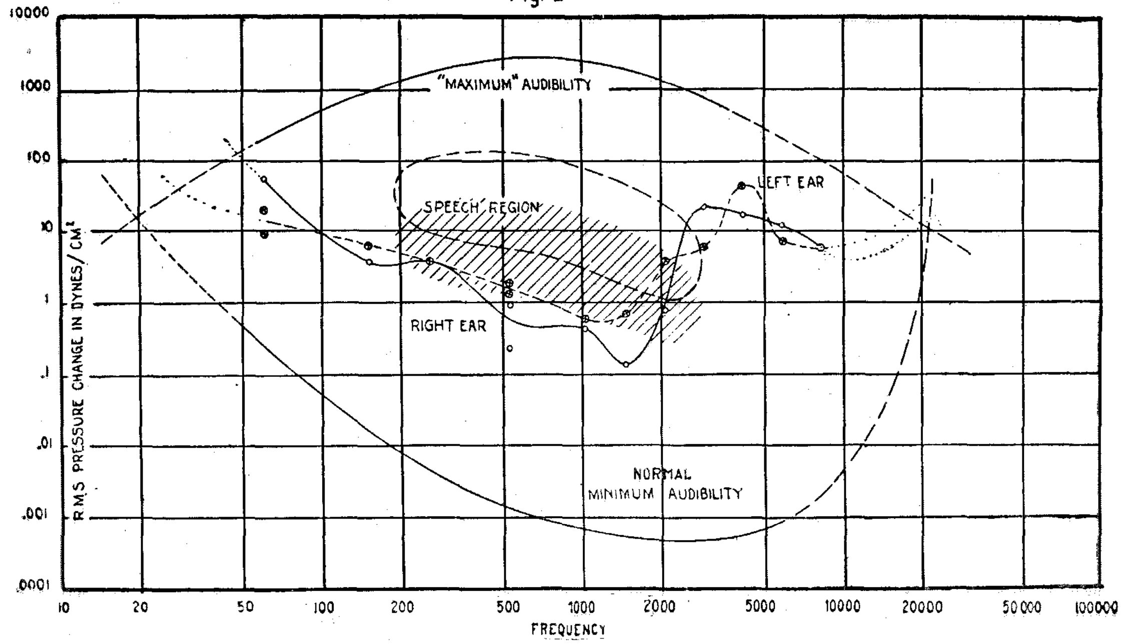
---

[2] The extrapolation upward of the curve of minimum audibility is consistent with some recent observations of Mr. C. E. Laine at the University of Iowa, *Physical Review*, May, 1922.

of the diaphragm and ossicles of the middle ear are probably so large that the nerves feeding these movable parts are stimulated. This observation at low frequencies as indicated in this work lends color to the hypothesis of otologists that abnormalities in the hearing of low frequencies are due to pathological conditions in the middle ear. This point is probably related to the tests on flexibility of the ear drum or ossicular chain due to the application of air pressure as observed by otologists in examination. Loss of sensitivity at low frequencies is considered an indication of obstructive deafness if there is no loss at high frequencies.

5. *Sensation Area.* From the combined standpoint of utility and logic the logarithmic relation between stimulus (pressure variation) and sensation can be assumed. The elliptical area between the two curves may then be taken to represent an area of sensation which is characteristic of the normal ear. Any point within this area represents a definite auditory sensation in frequency and intensity. The area of sensation is analogous to the field of vision of the eye. The part of this area which is most utilized in the interpretation of speech is represented approximately by the shaded area in Figs. 1 and 2 and corresponds in a way to the center of the field of vision. A normal listener tries, by keeping at a certain distance from a speaker, to bring this part of his sensation area into play in the same way that when examining an object he directs his eyes so that it falls in the center of the field of vision.

Fig. 2



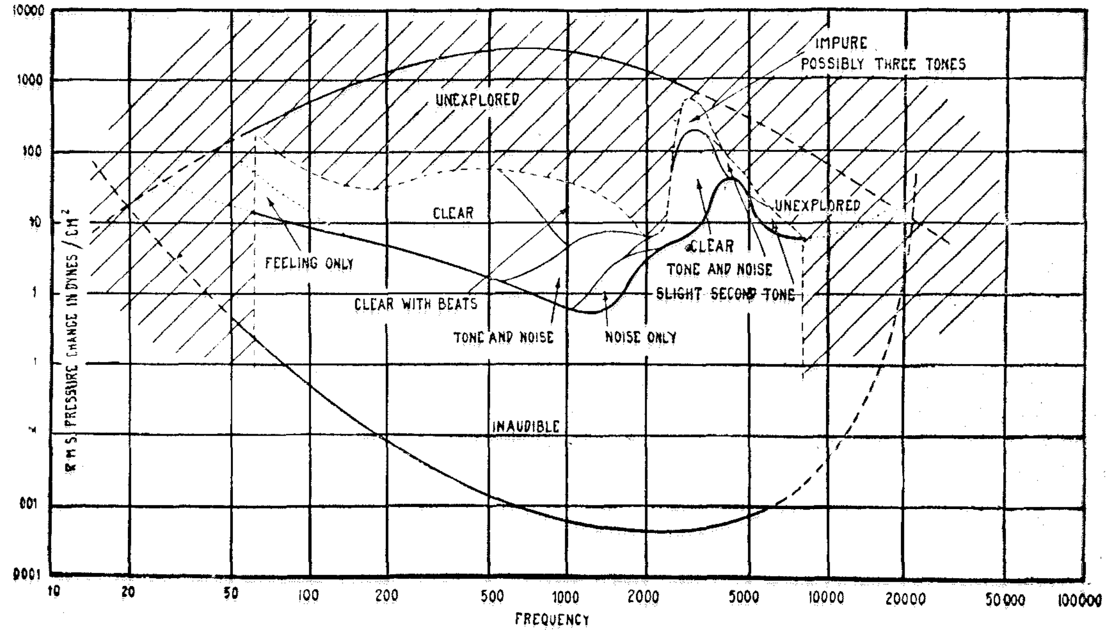
SENSATION AREAS (CHK)

An abnormal ear may be regarded as having an area of sensation which is smaller than the normal area but included within it. Fig. 2 is a plot of the minimum audibility of the right and left ears for a man (CHK) having a " catarrhal " deafness. The areas between these curves of minimum audibility and the curve of feeling are his areas of sensation. It will be seen that CHK retains about 50 or 60 per cent of the normal amount of sensation. He hears and interprets conversation with some difficulty.

Since the CHK curves pass through the speech region, part of it is entirely inaudible and the remainder is near minimum audibility for him. In order to make him hear well, the speech area must be raised to a higher level of intensity or loudness as indicated by the dotted curve.

In general it takes a loss of about 20 per cent of the sensation area to become noticeable and much more is disagreeable. A loss of 50 per cent requires the use of deaf apparatus. A loss of 75 per cent can be aided considerably by the use of high powered amplifying apparatus.

Fig. 3



QUALITY AREAS-LEFT EAR (C.H.K)

6. *Importance of Various Intensities in Speech.* It is interesting to speculate on how CHK interprets speech. It has been shown [3] that the intensity of speech may be varied over perhaps 70–80 per cent of the range of sensation without serious loss of intelligibility to the normal ear. As the sound intensity is decreased, the intelligibility drops very suddenly to zero at minimum audibility. A similar drop is to be expected at an intensity so loud as to be painful. It is evident, therefore, that the range in which speech is intelligible for CHK is very considerably limited as compared to normal. It is possible to design a deaf set which raises the intensity of the principal speech region to any desired place within the abnormal sensation area and so in a measure, compensate for this narrowed range. The region in Fig. 1 " Region of Lesser Importance in Speech," corresponds to stimuli in conversation of lesser energy content, such as the minor shadings and fainter consonant sounds. While it is physically possible to produce an amplification of speech so that this region is raised into the diminished area it is impracticable to do so because of the pain which would be caused by the louder components. A diminution in sensation area can, therefore, be only partially compensated for. In case the area is extremely narrow a deaf set furnishing optimum volume can only serve as an aid to lip reading.

7. *Quality of Hearing.* The sensation of a normal ear at any point in the auditory sense range (Fig. 1) may be described by a number of different adjectives, such for example as "clear," " musical," " even," " sustained," " smooth," " pure," etc. Such a description may, in fact, be taken as a reasonable indication that the quality of sensation at the point in question is normal. Abnormal ears sometimes experience a subjective degeneration of quality of pure stimuli which they describe as " rough," " harsh," " sharp," " buzzing," " vibrating," " hissing," etc. This subjective degeneration is independent of any tinnitus or head noises which the patient may have. Fig. 3 shows various regions of the sensation area which are degenerated in the case of CHK, left ear. The shaded area was not explored. The boundaries of the degenerated regions are usually more sharply marked than

---

[3] H. Fletcher, *Journal of the Franklin Institute*, June, 1922.

the outer boundaries of the sensation area. The sensations in these areas are so radically different from the sensation of a pure tone that it is with difficulty that the patient is convinced that the stimulation is the same pure tone to which he has been listening at the other intensities. The subject of these tests is a violinist and capable of better descriptions and finer distinctions than average.

Since all speech sounds may be considered as stimuli composed of various frequency components of certain intensities, the sensation caused by such a sound may be represented on this plot by points, or by a line provided the sound has a band spectrum. If the points or line falls within the sensation area the sound is audible. It is easy to see that if the points or the part of the line which represent those frequency components most essential to interpretation of the sound, fall within any of these abnormal areas, the sound is very likely to be misinterpreted. This adds a further source of loss in intelligibility to that already observed due to a narrowing of the sensation range. When an amplifying deaf set is designed, due care should be taken to raise the principle speech region in such a way as to cause a minimum overlapping with the abnormal areas.

Many practically normal ears have very small abnormal areas. They have always been found near minimum audibility and if this is always true would, therefore, have little influence on the hearing of the individual. They seem to be associated with "catarrhal" conditions, atlhough this cannot be stated positively.

8. *Binaural Sense.* The normal individual has learned to interpret the differential sensations of the two ears to advantage. It helps him to locate the direction from which sounds come, to have a sort of sense of orientation with respect to sounds approaching from different directions, and whether for physical or for purely psychological reasons to assist in focusing of the attention on one sound of a large number. Two ears also assist the individual in perceiving equally well sounds coming from different directions.

When one ear becomes less sensitive, even though the loss is small, the use of the binaural sense disappears and after a time is not missed, the subject depending upon other means of locating sounds. For the binaural sense to be most effectively utilized it is necessary that the ears be very nearly alike. When a binaural deaf set is made and fitted to a person with compensating sensitivity for the two ears so that both hear sounds equally loud, the sensation is usually so novel, that if the patient is actually able to experience a binaural sensation he is very much pleased. Usually, however, he has not used his binaural sense for so long a time that it takes a considerable amount of practice before he is able to have binaural experiences. It may be noted in this connection that the same experience is encountered in fitting the eyes with glasses. It is found that people with two eyes which are slightly different do not see stereoscopically, but if glasses are made so as to compensate and make the eyes nearly alike, it usually takes a certain amount of practice before the sense of perspective can be brought back.

APPENDIX

9. *Experimental Methods.* In order to discuss the principles of ear sensitivity measurement on a rigorous dynamical basis it will perhaps be clearer to describe briefly the experimental method used in producing known sound pressure in the ear canal at the various frequencies and intensities.[4]

As a source of sound, a small thermal receiver unit was used. This consisted of about twenty very small loops of Wollaston wire contained in a brass case small enough to be inserted in the external ear canal and entirely stop it up. In the average ear a volume of about 1 cm.^{3} of air is included between it and the drum membrane. A direct heating current is passed through the receiver and an alternating current of the desired frequency and intensity is superimposed to modulate its temperature. This modulation in temperature causes alternate expansion and contraction of a very thin film of air covering its surface and so produces alternations in pressure in the ear canal of the frequency of the impressed alternating current. The intensity is proportional to this alternating current if it is maintained small compared with the direct current. This arrangement permits of producing

---

[4] For further details, see "The Frequency Sensitivity of Normal Ears," H. Fletcher and R. L. Wegel, *Physical Review*, June, 1922.

alternating or sound pressure on the ear drum with a comparatively simple dynamical relation between the source of sound and the ear drum. The thermal receiver is also dynamically one of the simplest sources of sound known.

The sound or alternating pressure was determined by calibration. This was done by inserting the thermal receiver in an air cavity of 1 cm.^{3} volume in front of a condenser transmitter diaphragm by which the alternating pressure developed by a given current in the receiver could be measured.[5] By measurement of the current for minimum audibility or "maximum audibility" or for any other intensity the pressure in the ear canal is determined.

10. *Dynamical Principles of Ear Measurements.* From a dynamical standpoint the phrase "sensitivity of the ear" as it is usually used is rather indefinite. When a figure is given in ergs per second, the rate of flow of energy through an area equal to that of the ear opening in an unobstructed wave, is commonly meant. This has no simple relation, theoretically at any rate, to the net rate of flow of energy into the ear when the head is placed as an obstruction to the wave. The distortion of the sound field by the head varies greatly with frequency. Similarly, there is no simple relation between the energy flowing into the ear and that transmitted to and absorbed by the ear drum or by the cochlea. In the experiments recorded above, attention was paid to the experimental set-up so as to make the figures given have a more definite dynamical significance. Sensitivity is given in terms of the alternating (root mean square) pressure to produce a minimum audible sensation. The term "pressure" has so far been used in a rather loose sense. Just why this is so will be seen from the following argument.

The simplest method of describing the constants of a mechanical system is in terms of the components of its mechanical impedance and their relative dispositions in the same way that an electrical circuit is described by giving its resistance, inductance and capacity and the way in which they are connected. In a linear system having a single degree of freedom, the impedance may in general be written in the form

$$Z = r + wjm + s/jw.$$

The symbols are as follows;

$j = \sqrt{-1}$,

w = 2π times the frequency,

r = frictional resistance to motion, with respect to a stationary body and involves dissipation of energy at a rate of $\dot{x}^2 r$ where $\dot{x}$ is the root mean square value of the relative velocity. The velocity $\dot{x}$ will be assumed simply sinusoidal in what follows,

m = mass or inertia constant involving an average storage kinetic energy of $\dot{x}^2 m$ through one cycle,

s = stiffness constant involving an average storage of potential energy through one cycle of $\dot{x}^2 s/w^2$.

If the r. m. s. alternating force acting is F, the motion at any frequency is given by

$$\dot{x} = F/Z.$$

In analyzing a system in which the constants may be considered as "lumped," that is in which, for the purpose of practical solution, a finite number of degrees of freedom may be assumed, the method is to find the most useful way of "lumping" these constants. The motions are then represented by a series of equations, one for each degree of freedom, between the forces acting and the impedances and velocities. The determinant of the coefficients of these equations is the Lagrange determinant of the system. The only caution to be observed in lumping the constants is that the reciprocal relation, which is a property of any linear system holds also for the physical system which the assumed Lagrange determinant is supposed to represent.

The method may be illustrated by the following application to the sensitivity measurements described above.

The dynamical system used in calibration with the condenser transmitter consists of three parts;

(a) The very thin pulsating air film over the thermal receiver filaments. The expansion of air around the wires is represented by the "diffusion" equation, the solution of which in such a case of cylindrical symmetry is given as a Bessel's function of the distance from the

---

[5] For the method of calibration of the condenser transmitter, see article by H. D. Arnold and I. B. Crandall, *Physical Review*, July, 1917.

wire.[6] This wave is so quickly damped in travelling away from the wire as to be negligible beyond the first zero point of the Bessel's function. The vibrating system of this receiver may then be considered as a cushion of air next to the wire of a thickness a little less than the first half wave length of the heat wave. The thickness of this cushion is an inverse function of the frequency.

(b) The air chamber between the thermal receiver and condenser transmitter diaphragm having a volume of 1 cm.^{3} and enclosed by practically unyielding walls with no openings.

(c) The condenser transmitter diaphragm, being stretched very tightly and air damped. It may also be regarded as unyielding, or as having an impedance very high compared to that of the connecting air chamber.

If for simplicity the mass reaction and internal losses in the air chamber may be neglected, it may be seen that the moving system of the receiver may be regarded as a weightless and frictionless "diaphragm" surrounding the wires at a distance equal to the effective thickness of the active air film and may be shown to have an intrinsic stiffness reactance of:

$$Z_1 = \frac{s_1}{jw} = \frac{\gamma p_0 a_1^2}{jwv_1}.$$

In this expression, γ is the adiabatic constant of air, p_0 the atmospheric pressure, a_1 the area of the fictitious diaphragm, and v_1 the volume of air in the film. This diaphragm is loaded externally by the air chamber, when the transmitter diaphragm is prevented from moving, by a stiffness reactance of

$$M_1' = \frac{S_1}{jw} = \frac{\gamma p_0 a_1^2}{jwv},$$

in which v is the volume of the air chamber. Similarly, the load of the air chamber on the transmitter diaphragm, whose area is a_2, is

$$M_2' = \frac{S_2}{jw} = \frac{\gamma p_0 a_2^2}{jwv}.$$

The air chamber also acts as a mutual impedance between the thermal unit and the transmitter diaphragm equal to

$$M_{12}' = \frac{S_{12}}{jw} = \frac{\gamma p_0 a_1 a_2}{jwv}.$$

If, further, the intrinsic impedance of the transmitter diaphragm, which may be any function of frequency, be denoted by Z_2, the equations of motion of the system may be written

$$F = (Z_1 + M_1')\,\dot{x}_1 - M_{12}'\,\dot{x}_2,$$
$$0 = -\,M_{12}'\,\dot{x}_1 + (Z_2 + M_2')\,\dot{x}_2.$$

In these equations, F is the force acting on the thermal receiver "diaphragm" due to alternating current, $\dot{x}_1$ the velocity of its motion and $\dot{x}_2$, the velocity of motion of the condenser transmitter diaphragm.

A rough calculation shows that v_1 is very small compared with v, so that S_1 may be neglected compared to s_1 and that the reaction $M_{12}\,\dot{x}_2$ may be neglected. The analysis of the condenser transmitter shows Z_2 to be very large compared to M_2. These equations may then be rewritten

$$F = Z_1\,\dot{x}_1,$$
$$M_{12}'\,\dot{x}_1 = Z_2\,\dot{x}_2. \qquad (1)$$

The equations of motion, when the receiver is inserted in the ear, may be derived in a similar way. In this case, although the volume of air between the receiver and ear drum is the same as before, the walls may yield appreciably, particularly in some frequency ranges. The mutual impedance between the receiver and ear drum, is, therefore, not necessarily a simple stiffness reactance. Also the loads due to it on the thermal receiver and ear drum, which in this case takes the place of the transmitter diaphragm, are not simple stiffness reactances. The constants in the case of the ear system will be denoted by the same letters as those used in the calibration, but with the primes dropped, with the exception that the intrinsic impedance of the ear drum is denoted by D. D includes the reactions of the ossicles of the middle ear and the cochlea and is probably a complicated function of frequency. If, as may be expected, nature's design is efficient, then D must be of the same general order of magnitude as the load on the ear drum, M_{12}, of the ear canal. This probably constitutes the largest difference between the calibration and the observational systems. Strictly, of course, the condition for maximum power absorption by the ear drum from the air is that D be the conjugate of the impedance of the load on it due to the unobstructed ear canal. This condition is not obtained in nature because of such requirements

---

[6] See Wente, *Physical Review*, April, 1922.

placed on the design as protection from injury, etc.

In the case of the ear, M_1 may again be neglected, compared to Z_1, and the reactance, $M_{12}\dot{x}_2$ may be neglected. Then

$$F = Z_1 \dot{x}_1,$$

$$O = - M_{12} \dot{x}_1 + (D + M_2) \dot{x}_2, \qquad (2)$$

where $\dot{x}_2$ represents the velocity of motion of the ear drum. Suitable variations with frequency are implied in each of the " constants " of this system.

We are now in a position to see just what has been measured and called, for the sake of brevity or want of a better name, " minimum audible pressure " in the first part of this paper.

Let $\dot{x}_1$ now represent the velocity of the receiver diaphragm in both systems corresponding to that necessary to obtain a minimum audible sensation in the ear, and F the corresponding force. Then $\dot{x}_2$ will be the velocity of the ear drum corresponding to minimum audibility in equation (2). In the calibration, the pressure p' on the condenser transmitter diaphragm corresponds to $\dot{x}_1$. The total force acting on this diaphragm is $p'a'$ where now a' designates its area. Since this force is relieved by the motion of the diaphragm, it is seen from equation (1) to be equal to

$$p' a' = M'_{12} \dot{x}_1. \qquad (3)$$

Similarly if the actual pressure on the ear drum is p, and its effective area, a, the total force on the ear drum $pa = D \dot{x}_2$. Combining equations (2) and (3) gives

$$p' = \frac{a}{a'} \frac{M'}{M} \left(\frac{M_2 \dot{x}_2}{a} + p \right), \qquad (4)$$

or

$$p = p' \frac{M}{M'} \frac{a'}{a} - \frac{M_2 \dot{x}_2}{a}. \qquad (5)$$

The pressure p is the actual pressure on the ear drum. The pressure p' is that measured and plotted in the diagram. If the walls of the ear canal and the ear drum were unyielding, p and p' would be identical for then $M = M'$ and $M_2\dot{x}_2/a$ would vanish. If the yield of the ear canal walls were such as to relieve half the pressure in the canal and that of the ear drum about the same, the difference would be considerably less than one of the divisions, in the diagrams, on the intensity scale. If the drum impedance D should be found to be negligible compared to its load M_2 the difference would be considerable. This, however, is hardly to be expected even through narrow ranges of frequency. If the impedances in the formulas were measured the energy flow into the ear drum could be computed.

In conclusion, the present status of the ear problem may be summarized. The philosophy of external ear dynamics has been touched on but there still remain difficult problems both theoretical and experimental. A start has been made on a sound basis in the explanation of the action of the cochlea by Roaf, " Analysis of Sound Waves by the Cochlea," *Philosophical Magazine*, February, 1922. Nothing dependable has as yet been published on the action of the middle ear for audio frequencies. It is usually assumed that the various parts undergo relative displacements at audio frequencies in the same way as they react to static forces but this is very likely far from the truth.

The Binaural Location of Complex Sounds

By R. V. L. HARTLEY and THORNTON C. FRY, Ph.D.
Engineering Department, Western Electric Company

(NOTE: *Much has been written on the subject of the binaural location of pure tones but the case of complex sounds has received little attention in recent literature. The purpose of the present paper is to bring the discussion of complex sounds abreast of that relating to pure tones. Those who wish to acquaint themselves with the work on pure tones will be interested in reading the theoretical work of the authors and the experimental studies carried out by G. W. Stewart and students working under his direction. This work has been reported in various papers, most of which have appeared during recent years in the "Physical Review" and the "Physikalische Zeitschrift."*—EDITOR.)

THE need of determining the location of enemy submarines and aeroplanes during the war brought into use practical methods for locating a sound source which depend upon differences between the sound waves reaching the two ears. This stimulated a general study of the phenomena involved in binaural sound location. The foundation for this study had already been laid in the work of Lord Rayleigh and others, who, following more or less in his footsteps, had accumulated a considerable amount of information of both theoretical and experimental sorts. Of this information almost all that was of a theoretical nature and a considerable portion of the experimental kind dealt only with the location of *pure* tones, the more complicated and in some respects more important problem of *complex* sounds being almost entirely neglected. Such advances as were made in the theoretical aspects of the problem during the war were subject to the same restriction so that even to-day no comprehensive theory has been advanced which adequately covers the problem of the location of such sounds as occur in every-day life, and in the practical applications of binaural methods. However, the results obtained with pure tones can be made to throw considerable light upon the problem, and it is primarily from this standpoint that the following discussion is written.

It may be well at the outset to review some of the outstanding differences between the observed phenomena in the two cases. The accuracy of location is much less for pure tones, as is also the sense of definiteness of the sound image. The location of pure tones is almost wholly binaural as is evidenced by the inability of persons deaf in one ear to locate such a tone. With complex sounds not only is the location by binaural effects more accurate and definite, but also the observer is not dependent on these alone. Persons who are deaf in one ear can locate familiar complex sounds almost as well as those with normal hearing.

Practically all theories of sound location start from the assumption that the listener subconsciously observes certain sound characteristics which depend upon the position of the source and forms a judgment of where the source must be by comparing these characteristics with information which he has stored up as a result of his past experience with cases in which the position of the source was known. In order to fix the position of the source he must assign to it three coordinates such as its distance and some two angles which define its direction. To do this he must be able to observe at least three independent properties of the sound which are functions of the position of the source. If fewer than three are available some difficulty in location is certain to arise. If more than three are available there is the possibility of a number of simultaneous independent determinations of the three coordinates.

If the sounds of every-day life were never distorted in transmission all of these determinations would yield the same set of coordinates and the only advantage which the listener would gain from the additional information available would lie in the fact that some one set might be peculiarly sensitive to slight differences in the position of the source, and therefore might lead to increased certainty on the part of the observer. Owing to reflection from the walls of buildings and the like, the sounds of every-day life seldom arrive undistorted, so that the observer must always be somewhat uncertain as to whether or not the coordinates of the sound source are actually those which he deduces from the properties of the sound wave as it reaches his ears. If enough properties are available to permit him to make two independent determinations he may use one of them to check the other, and if they agree he is justified in a feeling of increased certainty as to the accuracy of his judgment. The more independent determinations he can make the more checks he will be able to apply and consequently the more confident he will be.[1]

It should not be inferred, however, that it is only the sounds of the street which reach the observer in a distorted form. In a great many laboratory experiments the characteristics of the sounds have been inconsistent, and in some cases they have not even corresponded to any actual source whatever. Under these circumstances, if an image is formed at all, some purely psychological factors must enter in. For pure tones it has been found possible to explain much of the experimental data obtained under circumstances such as this by assuming that the observer subconsciously judges one or more of the characteristics to be in error and applies such corrections as will make all of the data corresponds to an actual source. As a criterion for determining which characteristics will be altered, it is assumed that, in general, those are chosen which require the smallest changes.

Let us now consider what characteristics are available for locating sounds of different kinds. A pure tone from a source at rest with respect to the observer has at any point only two physical characteristics which are subject to change with the position of the source. They are its amplitude and phase. Corresponding to each position of the source there is a particular amplitude and phase at each of the two ears so that a total of four properties—the loudness of the sound, the average phase, the difference in amplitude (which may conveniently be expressed as a ratio) and the difference in phase at the two ears—are available for determining the position of the source. It is inconceivable that the average phase can have anything to do with the location of the sound since it may be changed at will without altering the position of the source. The same remark applies to the loudness of the sound except in those instances where the observer is familiar with the source to such an extent as to know how loud it may be expected to be. Hence, if we restrict ourselves to the cases in which prejudicial information of this sort does not exist, we find that the observer has only two quantities from which he may deduce the position of the source. We should therefore expect that these two quantities would make it possible to locate the tone with respect to two coordinates only. This is found to be in general agreement with experiment, for most observers locate all sources of pure tones in the same horizontal plane with their heads and determine only the distance and angular departure from the median plane. If the source is more than a few yards away the intensity ratio and phase difference change very slowly with distance so that in this case even the sense of distance is not keen and a feeling of certainty exists with respect to the direction only.

In many experiments the tones at the two ears have been varied arbitrarily so as to give combinations having equal phases and unequal intensities or vice versa—combinations which cannot arise from actual physical sources in the absence of distortion. Under these conditions the observer generally corrects one to a value consistent with the other except in extreme cases where the correction required for this purpose would be inordinately large. When this occurs he may either assume both to be correct and form two images—one based on the phase difference together with a mentally supplied intensity ratio consistent with it, and the other similarly derived from the observed intensity ratio—or he may fail to have a sense of location at all.

Before considering the available characteristics of complex sounds in general let us confine our attention for a time to those which are made up of a limited number of sustained pure tones such as an organ note with its series of overtones, or a group of tuning forks. Here the number of characteristics increases rapidly with the number of component tones. For each component tone there are two quantities: intensity ratio and phase difference. In addition, at either ear alone the relative intensities

---

[1] It is interesting to note in this connection that it is not surprising that an observer locates a complex tone with much greater certainty than a pure tone when we consider how rapidly the number of independent sets of data increases with increase in complexity of sound. We have already said that three independent properties are needed for the determination of the three coordinates of the source. Hence if only three are available, only one determination can be made and no checks are possible. On the other hand, if four are available, four groups of three each can be formed and therefore four separate determinations can be made. Similarly, 10 determinations can be made from 5 properties, 20 from 6, and 120 from 10.

of any two of the tones changes with the position of the source, owing to the diffraction of the sound waves around the head being different for different frequencies. There are therefore as many of these observable intensity ratios as there are pairs of components. Similarly, for any two tones whose frequencies are commensurable, the relative phases of the two at the same ear depend upon the position of the source.

Not all of these characteristics are capable of contributing to binaural as distinct from monaural location. In fact, only the phase differences and intensity ratios of the separate components are binaural. A man who is deaf in one ear has available all of the relations between the intensities and phases of the various components at his normal ear. That these relations do actually contribute to sound location is supported by experimental evidence. Myers [2] found that, after familiarizing himself with a complex sound, a blindfolded observer could locate its position with considerable accuracy, even when it was moved about in the median plane, but that his accuracy could be destroyed by varying the relative intensities of the components.[3] It is not surprising, then, that for complex sounds the accuracy is about the same whether the location is binaural or monaural.[4] The observed failure of monaural location in the case of a pure tone follows directly from the absence of other frequencies with which the pure tone may be compared.

As we are here concerned with binaural phenomena we shall confine our attention to the relative phases and intensities at the two ears. The question at once arises: does the observer actually hear the different tones separately, and if so, does he assign a location to each separately?

To what extent the listener locates each component separately depends upon the ease with which the tones can be distinguished. The experiments which bear most directly upon this point are those in which the component tones at the two ears are arbitrarily adjusted to give values of phase difference corresponding to different locations. This is done under conditions where the location of each component separately is largely determined by the phase difference. More[5] experimented with two tones, transmitting them to the ears through tubes of adjustable lengths. This permitted him to change the phase difference at the two ears while keeping the intensities substantially equal. He observed the apparent location for various settings when each tone was applied by itself and when both were applied together, using forks of 256 and 320 cycles. With the paths equal the tones combined into a chord located in the median plane and the separate components could not be heard. With a setting for which the two components separately appeared on opposite sides of the head, one component was heard distinctly by the right ear only on the right side, and the other by the left ear only on the left side. At the same time the chord was heard rather indistinctly near the median plane but tending slightly toward the side of the lower tone.

Apparently the observer does not consciously separate the chord into its components unless he is forced to do so by some inordinate discrepancy between the positions of the images formed from them. There is no evidence in the case of equal paths to show that he did or did not subconsciously locate the separate components and find them to be in agreement. In view of the second experiment it seems probable that he did. In this latter experiment he obviously found that the two components corresponded to different locations and assigned different sources to each. At the same time his experience told him that tones which would combine to form a musical sound generally have a common source. Hence he may have concluded subconsciously that the sound waves had probably been distorted in coming from a common source and so he corrected his observations on both tones to make them consistent and arrived at an image of the chord between the other two.

Similar results were obtained with forks of 256 and 384 cycles per second, except that in

---

[2] C. S. Myers, *Proc. Royal Soc.*, 1914, B 88,267.

[3] It should be noticed that this effect must have been purely psychological since it could be produced without moving the source at all. It therefore lends plausibility to the assumption upon which our theory is based: that when discordant or unusual stimuli are experienced, a mental readjustment of the stimuli is made in order to render them more nearly consistent with every-day experience.

[4] As shown by the experiments of Angell and Fite upon persons deaf in one ear. *Psychol. Rev.*, vol. 8, pp. 225–246, 1911.

[5] Louis T. More: *Phil. Mag.* XVIII, 1909, p. 308.

general the lower tone was completely blotted out. The higher tone was usually quite distinct and definitely located. The image of the chord was nearer to the image formed when the higher component was sounded by itself than to the image formed from the lower one alone. With settings for which the directions of the tones separately were the same, whether right, left, or middle, the upper tone disappeared leaving only the chord. In experiments with forks of 256 and 512 cycles it was difficult to distinguish the separate notes. With settings for which the two separately were on opposite sides the combination was on the side of the lower fork. This can be interpreted as meaning that the octave relationship is inherently difficult to resolve, or else that tones an octave apart so generally come from a common source that the observer was unwilling to make any other assumption.

Although the explanation of these results is not yet thoroughly understood, they show very definitely that in locating complex sounds made up of pure tones the observer does within limits locate the components separately. If they agree, a single image is formed; if they do not, he may either locate the tones separately or form a single compromise image or do both.

It is in this way that the theory developed for pure tones is applied to complex sounds made up of pure tones. The next step is to extend it so as to include complex sounds in general. To do this we must picture the observer as resolving each sound into sinusoidal components locating the components separately and forming one or more images based on a combination of the apparent sources as indicated by the separate components. While it is fairly easy to effect such a resolution mathematically it is somewhat less easy to interpret the result in a manner satisfactory to our intuitive conceptions of the phenomena involved; also, granted the theoretical possibility of the resolution, there remains the question of what physical or psychological limitations there may be to its application.

In view of the fact that a really pure component tone has no beginning or end, and no fluctuations in its amplitude, it is not at once apparent how a single discrete sound such as the bark of a dog can be resolved into components of that nature. However, if enough components are available it has been established beyond question that by properly choosing their frequencies, amplitudes, and phases, a combination may be arrived at in which the algebraic sum of all the components is zero for all instants before and after the period occupied by the sound and equal to the instantaneous value of the sound wave for instants within that period. This combination is known to mathematicians as the Fourier Integral corresponding to the wave, and the formula for the phase and amplitude of each component sinusoid is known. It is an extension of the well-known Fourier series expansion used for resolving sustained periodic disturbances.

The physical interpretation of this integral may be facilitated by reviewing the steps in its evolution from the Fourier series. It is well known that if the sound in question were repeated at regular intervals the resulting periodic wave could be resolved by Fourier analysis into a series of sinusoidal components, the frequencies of all of which are integral multiples of the frequency of repetition of the sound. Successive components therefore differ in frequency by an amount equal to this frequency of repetition. Now it is not essential that the repetitions of the sound follow each other immediately. Instead, they may be separated by intervals of silence. The effect of such silent intervals is to reduce the frequency of repetition and therefore also the fundamental frequency. As a result the component frequencies are brought closer together and the number within any particular frequency range is increased.

Suppose now that the interval between repetitions is indefinitely increased. As this is done the effect of any one occurrence of the sound becomes more and more independent of the others, and in the limit when the sounds next preceding and next following the one under consideration are infinitely far removed, we have the case of a discrete sound. As this limiting case is approached the fundamental frequency becomes smaller and smaller and the component frequencies, which are multiples of it, are separated by infinitesimal frequency differences. While the amplitude of each component also decreases, the number of components increases at such a rate that the aggregate energy of all the components within a given frequency range remains finite. In this way, the distribution of the sound energy over various frequencies—that is, the "energy spectrum"—can be obtained.

It is evident, then, that when an aperiodic complex sound is resolved mathematically there results an infinity of component tones, each having a characteristic intensity and phase. If an observer were capable of an equally complete resolution he would have at his disposal an infinity of sets of data from which an infinity of images could be formed. In the absence of distortion these should all coincide.

Practically, of course, no such refinement of resolution is possible. The ability to distinguish differences in pitch varies from person to person, but the minimum intervals employed in musical composition probably give a rough measure of the normal resolving power of the ear. Even with this limitation the broad sound spectrum, such as an irregular sound produces, is capable of yielding a very large number of separable components; and hence a large number of individual images. It is this fact—that with a very complex sound the number of independent determinations of the image is limited only by the resolving power of the observer—which makes his accuracy of binaural location as well as his sense of certainty much greater for such sounds than for pure tones.

So long as the images of all the components coincide, it is of little importance how fine the resolution is, for further refinement only serves to increase the sense of certainty by adding to the volume of accordant evidence. However, when the images are not in agreement the problem is more complicated and the degree of resolution becomes important. Here also purely physical considerations cease to be adequate and psychological factors must be considered similar to those involved in the location of a pure tone for which the intensity ratio and phase of difference do not correspond to any actual source. When an observer is faced with discordant results he must make some subconscious judgment. For small discrepancies such as occur in every-day experience, he probably assumes those images which depart most from the rest to be misplaced because of distortion during transmission and so either corrects or ignores them. If the discrepancies are large he may find it difficult on the ground of experience to believe that so much distortion could occur. In such an event he will most likely form several images from different components or in extreme cases lose the sense of location altogether.

Bowlker found separate images to occur experimentally both for band music, which approaches a collection of tones and for the irregular barking of dogs. He placed tubes of unequal length to his two ears thereby upsetting the normal diffraction around the head and interposing a longer path on one side than on the other. Obviously, the distortion produced in this manner is of a type not likely to be met in every-day life and affects different frequencies in widely different fashions. He reports that when listening to "a band of three or four instruments played in the open—the notes will be found to be scattered over a wide range, most being to the side of the short tube, some being in front and some being to the side of the long tube. In listening with such a pair of tubes to two dogs furiously barking the effect is at first quite alarming—one seems to be in the middle of a pack of dogs some of which are rushing viciously at one's throat."

An illustration of failure to form any image is found in a phenomenon observed in the use of binaural compensators for determining the direction of submarine sounds. The sound is picked up by two submarine telephone transmitters and led to the ears through independent paths. By adjusting the lengths of the paths the image can be shifted from side to side and for practical purposes the setting of the instrument is made by bringing the image exactly to the middle. A fairly definite sound image is formed, but observers report that part of the sound does not merge into this sound image and move in response to the adjustment, but instead appears as a diffuse background of noise.[6] This may be explained on the assumption that, while the images formed from most of the sound components agree sufficiently well that the observer corrects them to a single position, certain components are so distorted by resonance effects inherent in the apparatus that their images are scattered more or less at random. The lack of agreement among any con-

---

[6] This interesting phenomenon was called to our attention by Mr. Richard D. Fay of the Submarine Signaling Corporation who tells us that it has been noted by a large number of observers.

siderable number of these prevents the formation of a second image and causes the sense of diffusedness.

As the distortion becomes still more extreme we should expect the experimental results to depend more and more upon the observer's power of resolution, for as the distortion is progressively increased a condition must finally be reached where the positions of the images are appreciably different for two components whose frequencies are so nearly alike as to make their recognition as separate tones difficult if not impossible. This condition actually occurred in an experiment of Baley's with a sound consisting of a mixture of sustained tones. Its effect on the listener is interesting from the standpoint of subconscious readjustment of discordant data.

Baley's[7] experiment consisted in applying a number of sustained tones to one ear of a muscially trained observer and a number of different tones to the other ear, and testing his ability to assign them to their proper sides. So long as the intervals between the tones were fairly large, the observer never failed to locate them correctly. Considering the entire stimulus as a complex sound we may think of the observer as locating the tones individually and finding them to fall definitely into two groups whose images are located one at each ear. However, when he used six tones which were separated from each other by a single tone interval, the separate components could not be distinguished and a painful sensation was produced. The observer was apparently faced with the situation that to make the observed intensity ratios and phase differences correspond to a single source would involve extremely large corrections in the observed data. On the other hand, his power of tone resolution was insufficient to separate the components and assign them to different sources. It is not surprising, then, that the difficulty manifested itself by painful sensations. While this illustration is taken from an extreme condition of laboratory experiment and may appear to have little bearing on the every-day location of sounds, it is really significant because of the manner in which it illustrates the importance of psychological factors in all cases in which the sound waves are distorted.

RESUMÉ

In the foregoing discussion an attempt has been made to bring out the main features involved in extending the theory of the binaural location of pure tones to cover, qualitatively at least, the location of complex sounds. It has virtually been assumed that the latter involves three processes: first, the resolution of the sound into its component tones; second, the independent (generally subconscious) location of each separate component; and third, the formation of a conscious judgment of the position of the source based on the locations of the individual images. The greatly increased amount of data available when the sound is complex has quite different effects on the final result according as the different images do or do not coincide. If they do, the accuracy of location and the sense of certainty are increased. If they do not, confusion arises, subconscious corrections are called for, and the final result is likely to depend very considerably on the psychological processes and individual prejudices of the particular observer.

---

[7] Stephan Baley: *Zeit. f. Phychol. u. Physiol.*, v. 70, 1914, p. 347.

A Low Voltage Cathode Ray Oscillograph[1]

By J. B. JOHNSON, Ph.D.

Engineering Department, Western Electric Company

A CATHODE ray oscillograph tube operating at a comparatively low voltage was described by the writer some time ago before the American Physical Society.[2] Since then, the tube has been further improved and its operation studied so that now both the structure of the tube and the principles which have made the construction possible can be described in greater detail.

In the older types of Braun tubes the electron stream is produced by a high voltage discharge through the residual gas in the tube. This requires a source of steady potential of from 10,000 to 50,000 volts, an installation which is expensive, non-portable, and dangerous. In the new type of tube the low voltage operation has been obtained by the use of a Wehnelt cathode as the source of electrons, so that the lower limit of voltage is set by the effect of the electrons on the fluorescent screen and not by the voltage needed to obtain the electrons. At 300 volts the electrons produce quite bright fluorescence on the screen and the tubes are therefore designed to operate at 300 to 400 volts.

The external appearance of the tube is shown in Fig. 1. The electrodes are located at one end of the pear-shaped bulb, and the fluorescent material is deposited on the inside of the larger, flattened end. The tube is provided with a base which fits into a bayonet socket such as is used for vacuum tubes, and all the connections are made through the base. There are two orthogonal pairs of deflector plates inside the tube for electrostatic deflection, while magnetic deflection is produced by applying a field from the outside.

The internal structure differs considerably from that of previous forms of Braun tube and it will therefore be described somewhat fully.

THE FOCUSING

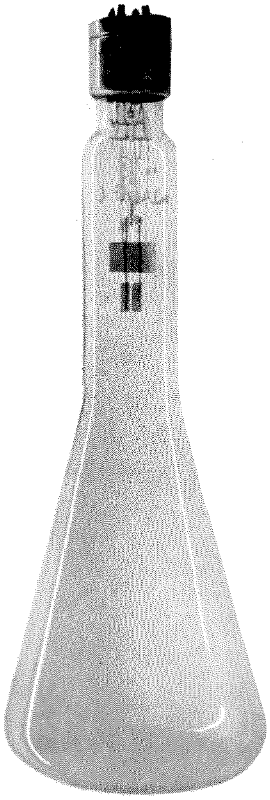


FIGURE 1

In some forms of Braun tube a sharp spot has been secured by using a very high voltage, and therefore high electron velocity, so that after the electrons have passed through one or two fine apertures to make the beam parallel there is not time enough for the mutual repulsion to spread the beam again appreciably before the electrons strike the screen. With other tubes an external "striction" coil has been used which maintains a strong longitudinal magnetic field in the region between the anode and the cathode and which brings the electrons to a focus on the screen. In the low voltage tube the spreading of the electron stream is greater than in high voltage tubes because of the greater time during which the mutual repulsion of the electrons acts, so that some means of focusing must be used. The electrons can be brought to a focus by a longitudinal magnetic field so adjusted that

---

[1] Also published in the *Journal of the Optical Society of America and Review of Scientific Instruments*, September, 1922.

[2] *Phys. Rev.* (2), Vol. 17, p. 420, 1921.

each divergent electron makes very nearly one complete turn of a spiral and in travelling the length of the tube returns to the axis at the screen. In this way a very sharp spot can be produced, but the sensitivity of the beam to deflection is reduced very much by the directing magnetic field.

The method of focusing that is used in the present tubes grew out of the suggestion by Dr. H. J. van der Bijl, that a small amount of gas be introduced into the tube. This gas, at a pressure of a few thousandths of a millimeter of mercury, serves to reduce to 1 mm. diameter a spot which would be 1 cm. across in a high vacuum tube. The sharpness of the spot depends also upon the current in the electron stream so that the focus may be controlled by the cathode temperature. The mechanism of this focusing action will be explained later.

The presence of this slightly ionized gas also serves the purpose of preventing the accumulation of charges on the glass, and it provides for the discharging of the fluorescent screen so that the electrons can drift back to the metallic circuit.

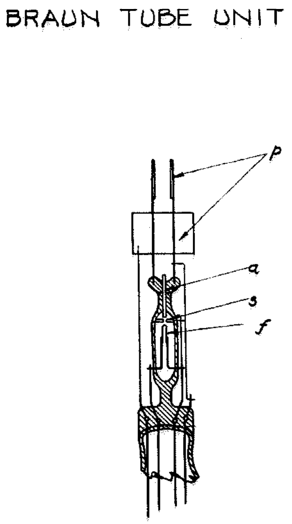


Figure 2

The Electrode Unit

With gas present in the tube, steps have to be taken to guard against arcing and the injurious effects of positive ion bombardment on the cathode. This is done by making the volume of gas surrounding the electrodes very small. For this purpose the cathode and anode, themselves small, are sealed into a short and narrow glass tube so that the volume exposed to both electrodes in common is less than 1 cu. cm. All paths between the electrodes are then so short that at this low pressure there is not sufficient ionization to build up an arc.

The structure of this unit, or "electron gun" is shown in Fig. 2. The cathode, *f*, is an oxide coated platinum ribbon of the same kind as the filament in our audion tubes. The anode, *a*, is a thin platinum tube 1 cm. long and 1 mm. in diameter, one end of which is about 1 mm. from the top of the filament loop, the other end opening into the main tube towards the fluorescent screen. Between the cathode and the anode and connected to the cathode is a metal shield, *s*, with a small aperture through which the electrons must pass in going to the anode. Nearly all of the electrons must then go to the inside of the tubular anode, and a small fraction of them pass through the whole length of the anode and form the beam in the main part of the tube.

The deflector plates, *p*, are also mounted rigidly on this unit. In order to avoid large differences of potential in the tube, one plate from each pair is permanently connected to the anode, the variable potentials being applied to the other plates. The complete unit is mounted at the small end of the tube with the anode and deflector plates toward the fluorescent screen.

The Filament

In some early forms the filament was bent into a simple hair pin loop which was placed close to the aperture in the shield. It was then found that the positive ions striking the filament from the direction of the anode soon destroyed the oxide coating and left the filament inactive. This trouble was largely overcome by placing the filament out of the direct path of the positive ions. The flat filament is now shaped into a ring as shown in Fig. 3, slightly larger in diameter than the aperture in the shield and is placed coaxial with the anode. The momentum of the positive ions then carries them past the active part of the filament

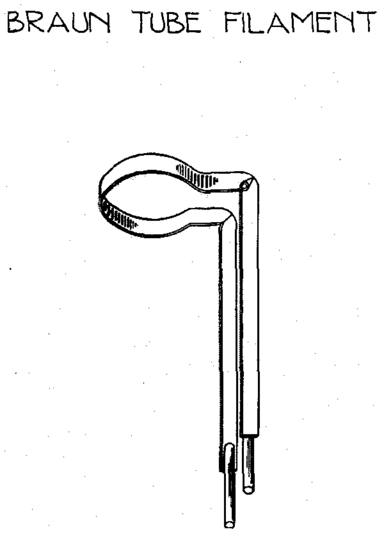


FIGURE 3

and they strike where little damage can be done. The length of service of the tube is still limited by the filament life, but this has been increased by the above artifice so that the tube now gives around 200 hours of actual operation.

THE DEFLECTOR ELEMENTS

The deflector plates are made of German silver, which is non-magnetic and which has a high specific resistance that diminishes the effect of eddy currents when magnetic deflection is used. The plates are 13.7 mm. long in the direction of the tube axis and the separation between them is 4.7 mm.

The sensitivity of the tube is such that the deflection of the spot is about 1 mm. per volt applied between the deflector plates. When using magnetic deflection, a pair of coils 4 cm. in diameter placed on the sides of the tube at the level of the deflector plates produces a deflection of approximately 1 mm. per ampere-turn flowing in the coils.

The electrons striking the screen drift back to the anode structure, and most of them are collected by the deflector plates. There is also a small ionization current flowing to the plates. The tube is therefore not strictly an electrostatic device, and this must be kept in mind when using it. Fig. 4 shows the current flowing to the two free plates at various voltages with respect to the anode. With the large positive values of plate voltage the current to the plates is practically equal to the current in the electron stream and consists largely of the returning electrons. The small current in the other direction when the plate voltage is negative is a measure of the ionization in the tube.

THE FLUORESCENT SCREEN

The screen is spread on the inner surface of the large end of the tube, using pure water glass for binder. The active material consists of equal parts of calcium tungstate and zinc silicate, both specially prepared for fluorescence. This mixture produces a generally more useful screen than either constituent alone. The pure tungstate gives a deep blue light which is about 30 times as active on the photographic plate as the yellow-green light of the silicate, while the silicate gives a light which is many times brighter visually than that from the tungstate. By mixing the two materials in equal parts a screen is produced which is more than half as bright visually as pure zinc silicate and more than half as active photographically as pure calcium tungstate.

For mechanical strength the end of the bulb which carries the screen is rounded outwards so that the screen is not a plane surface. This introduces a distortion of the fluorescent pattern which in most instances is negligible. If the pattern is recorded by a camera whose lens is D cm. from the end of the tube, then the apparent reduction of the deflection produced by the curvature of the bulb is given in terms of the deflection y approximately by

$$\Delta y = \frac{20 + D}{400\, D}\, y^3 \; cm.$$

THE FUNCTION OF THE GAS

The part which the gas plays in focusing the beam of electrons is an interesting phenomenon which depends upon the difference in the mobilities of electrons and positive ions. The electrons of the beam are pulled toward the common axis by a radial electric field produced by an excess of positive electricity in the electron stream and an excess of negative electricity in the space outside the beam. This distribution is produced as follows: Some of the electrons of the stream, in passing through the gas, collide with gas molecules and ionize them. Both the

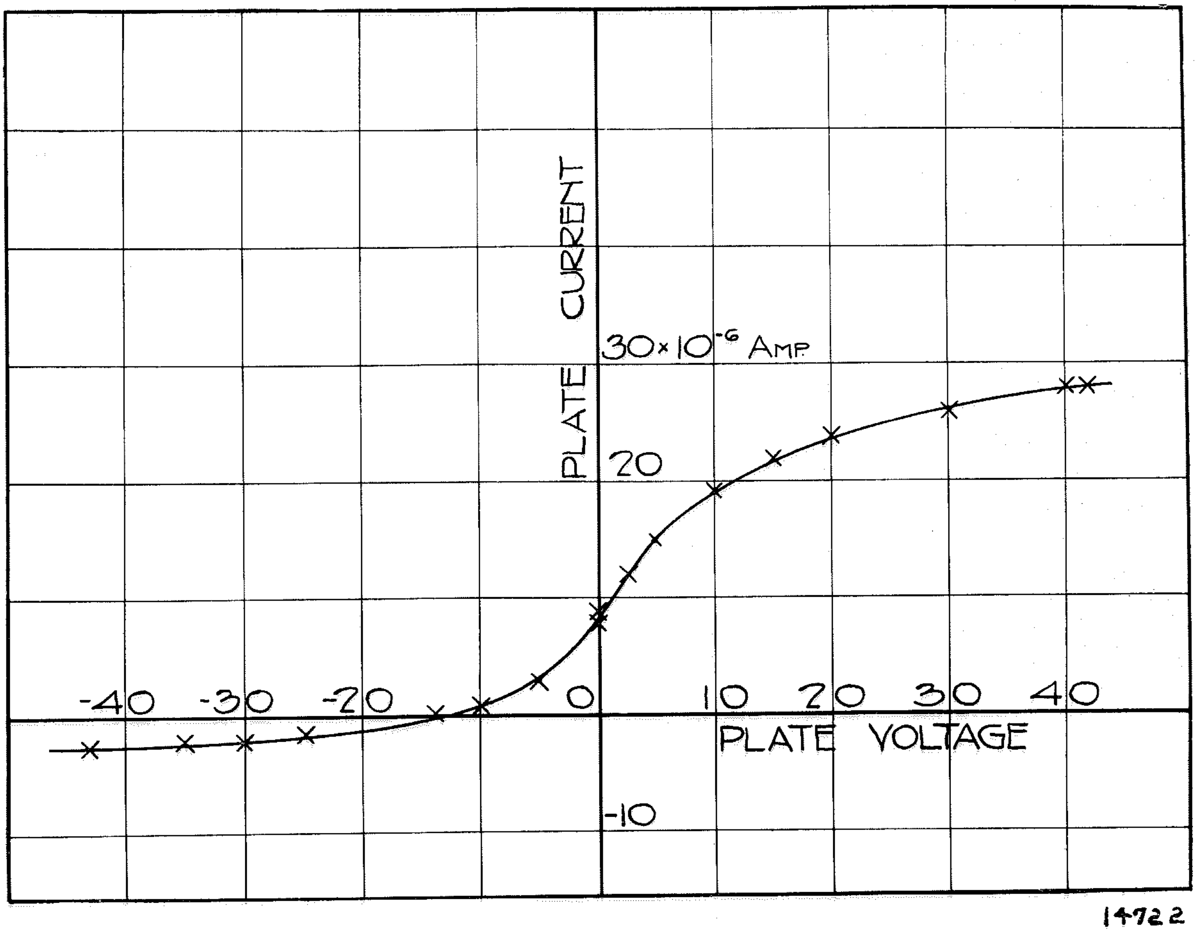


FIGURE 4

colliding electrons and the secondary electrons leave the beam but the heavy positive ions receive very little velocity from the impact and drift out of the beam with only their comparatively low thermal velocity. Positive ions, therefore, accumulate down the length of the stream and may exceed in number the negative charges passing along. At the same time, electrons are moving at random outside the stream, producing negative electrification. There is then a field surrounding the stream which tends to pull the electrons inward. If there were only the mutual repulsion between the electrons to compensate for, this would be done when the number of positive ions in the beam equals the number of electrons. There is in addition an original divergence of the beam which must be overcome. If this divergence is assumed to be one degree from the axis and the electron current 2×10^{-5} amp., then a simple calculation shows that the radial field required to pull the beam to a focus at the usual distance is about one volt per cm. This field strength is produced, with beams of the ordinary intensity, if there are four positive ions for each electron in the stream, a condition which seems not unreasonable.

The number of ions per electron in the stream is probably constant as the current in the stream is varied, since the conditions of collision and recombination are not altered. When the current is increased, therefore, the total positive ionization of the beam increases, the field around the beam becomes stronger, and the electrons are brought to a focus in a shorter distance.

These deductions have been confirmed experimentally. That the focusing of the stream depends upon the current flowing was one of the earliest observations made in developing the tube and this method has been used ever since to obtain a sharp spot. The point of convergence can be seen moving in the manner expected when the current is changed, and the effect has been further verified by using a tube with a movable fluorescent screen so that the length of the electron beam could be varied. The presence of the electric field around the beam was shown by the effect of two beams on each other, in a tube in which there were two electron streams crossing each other at right angles at their mid-points, each falling on a fluorescent screen. When one beam was moved away from the other by a field between the deflector plates, the second beam moved as if attracted by the first. The directed electrons in each beam were attracted toward the positive ionization in the other, and for one particular adjustment of the tube the displacement was such as would have been caused by a field of about 3 volts per cm., a result not far different from that previously calculated.

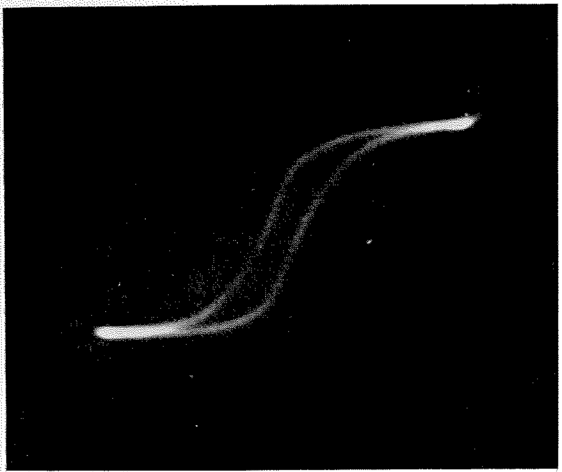


Figure 5

Since the beam must produce its own positive ionization some time must elapse before it can produce by collisions the required number of positive ions. Calculation shows this time to be of the order of 10^{-6} second. When the beam moves it has to build up the ionization as it goes along, and we should expect that when deflected very rapidly it might no longer be focused, due to lack of positive ions in its path. A test was made of this by applying a high frequency potential on the deflector plates so that the spot described an elliptic pattern. At a frequency of 10^5 cycles per second the line was still sharp, but at 10^6 cycles there was a noticeable widening of the line which is probably to be ascribed to imperfect focusing at this high speed.

In these experiments the evidence all points to the view that the focusing of the electrons is caused by an excess of positive charge in the beam itself, produced by ionizing collisions of the electrons with the gas molecules. Further confirmation is found in the fact that a focus is much more readily obtained in the heavier gases having slow molecules, such as nitrogen, argon or mercury vapor, than in hydrogen and helium where the mean velocity of the molecules is greater. The tubes are therefore filled with argon, the heaviest available permanent gas which does not attack the electrodes. The best pressure for the length of tube adopted and for the current which can be obtained in the beam is 5 to 10 microns, and this leaves considerable latitude for the adjustment of the electron current to get a sharp focus.

Examples of the Use of the Tube

Because of the small amount of auxiliary apparatus required with this form of Braun tube it has proved to be a very convenient laboratory instrument. It has found application in studying the behavior of vacuum tubes and amplifier and oscillator circuits, of gas discharge tubes, of relays, and of numerous other kinds of apparatus, both at low and at high frequencies. Some reproductions of photographs of various types of curves are given below to illustrate

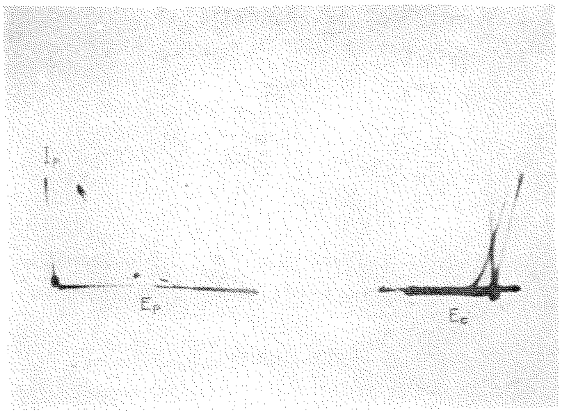


Figure 6

the kind of results which are possible with this oscillograph.

Fig. 5 shows the hysteresis curve of a sample of iron wire. The wire was placed in a small solenoid with one end toward the side of the tube. The magnetizing current passed through a resistance, the voltage drop of which was applied to one pair of deflector plates so as to give a deflection proportional to the magnetizing field. The stray magnetic field from the iron itself produced the deflection proportional to the induction. Alternating current was used, and the exposure was 20 seconds with lens opening *f* 6.3 and speed roll film.

In Figs. 6a and 6b are shown the current-voltage relations of an oscillating vacuum tube. The axes were obtained by grounding one or the other deflector element.

The measurement of modulation in a radio transmitting set has been reduced to a fairly simple process by means of the cathode ray tube. The low frequency modulating voltage, controlled by the voice, is applied to one pair of deflector plates, while the radio frequency output, with amplitude varying according to the low frequency voltage, is applied to the other pair of deflector plates. The resulting pattern on the screen is a quadrilateral of solid fluorescence, since the two frequencies are not commensurate. The two vertical sides indicate the greatest and the least amplitude of the high frequency, while the other two sides show the current-voltage characteristic of the transmitter. Fig. 7 shows such a pattern (retouched), the edges being much brighter than the centre. The exposure was two minutes using a Seed 23 plate and *f* 6.8 lens opening.

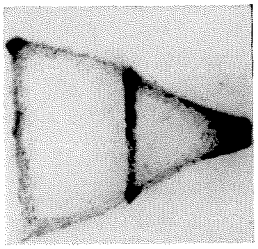


FIGURE 7

The Business Cycle. The Interest of the Telephone Industry in its Analysis

By F. E. RICHTER

Chief Statistician's Division, American Telephone and Telegraph Company

TEN or fifteen years ago it was the fashion among economists, in discussing the ups and downs of business, to concentrate their attention on crises and panics, and to endeavor to explain the causes for these recurrent disasters in the business world. Gradually the discussion broadened, and in more recent years there has been an increasing amount of study of the characteristics of each phase of what has come to be known as the business cycle. Causes and effects, inter-relations of forces, sequences of phenomena, and interpretations of all of these have been the subjects of serious attention and analysis on the part of economists and business statisticians. Neither students nor business men are by any means unanimously agreed on the many points of interpretation raised by such study; indeed, there is right at the outset a fundamental difference of opinion as to the existence of the business cycle itself as a distinguishable and definable economic phenomenon. There are those who consider the business cycle as being quite as inevitable, even if not quite as regular in recurrence, as the tides. There are others who would deny any validity to a theory of business cycles, or who refuse to see in the course of business any such sequences of events as can properly be given a name with the connotation of the word "cycle."

Whatever the degree of inevitability of a cyclical movement in business, the fact remains that business does not remain and does not tend to remain for any considerable length of time at a dead level of activity; apart from the natural tendency for the volume of production and trade to grow with the increase of population and with added human wants for economic goods, business as a whole has its ups and downs just as individual business enterprises have their periods of prosperity and adversity or setbacks. It is further true that certain types of phenomena tend to recur and to be sufficiently similar in their manifestations so that in spite of some differences in causes and in effects, it seems not unfair to describe the whole series of developments as a cyclical movement.

THE FORCES OF THE BUSINESS CYCLE

The chart, "The Forces of the Business Cycle" (Fig. 1), aims to depict what might be called a contour of business in the various phases of one of these cyclical movements, an imaginary one, to be sure, but roughly typical of such movements in general. The major phases of the cycle are shown on the chart as revival, prosperity, liquidation and depression. Every business man, whatever his views about business cycle theories, will recognize that these words describe the successive stages through which business now and again passes. If the physical volume of business grew merely with the increase in population, or with a regular increase in consuming power on the part of inhabitants of a country or of the world, it might be said to have a "normal" rate of growth or to be always "normal." If conditions of production also remained the same, or rather if an increased volume of production just sufficient to meet increased consuming power were possible and were not attended by changing costs or readjustments of any sort, business would lose much of its attractiveness as a venture and much of the dynamics and hazards which now characterize it. It is just because business is a struggle against varying natural and human forces and because business men and all other members of economic society are not equally able to forecast or cope with these forces, or are not equally affected by them, that the ups and downs of business come; in other words, that the business cycle develops.

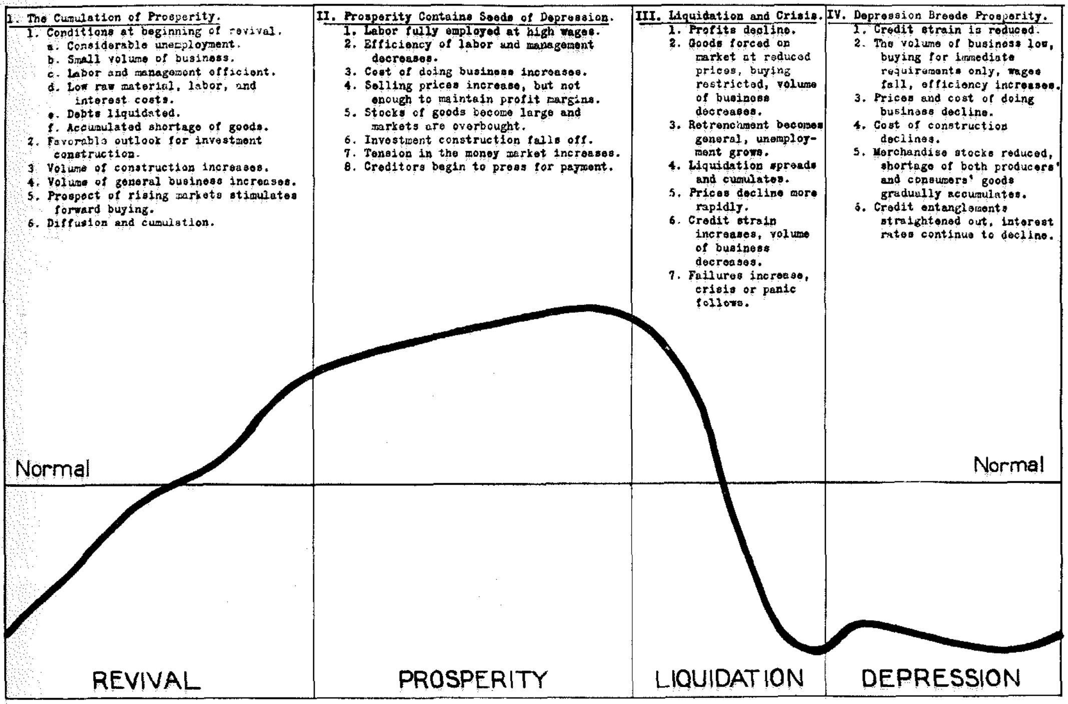
Since, therefore, business tends to fluctuate above and below what might be considered a normal rate of growth, the significance of the term "business cycle" has to do solely with the disparities between conditions of business at given times and what might be conceived to be normal condition of business at those times. The cyclical movement, further, is something quite different from and independent of the seasonal variations in business in general, or in any particular line of industry. There are few, if any, businesses which are not subject to seasonal fluctuations, and when the business statistician refers to a business cycle, he abstracts from these merely seasonal changes and takes them for granted, in the first place, just as he takes for granted a normal rate of growth.

The business cycle may be said to be due to the fact that apart from special influences, such as "lean" and "fat" years in agriculture, earthquakes, or other natural and humanly uncontrollable events, on the one hand, and wars or other political developments or events like strikes, on the other hand, it is apparently impossible for any mechanism as complicated as the modern industrial organization to function at all times and in all places with equal efficacy and smoothness; the economic forces which act upon the various parts of the huge industrial machine cannot and do not act with the same promptness and efficiency. This is true both for business as a whole and for individual industries within the whole.

THE PHASES OF THE BUSINESS CYCLE

Take, for example, the period of revival. It has been prepared for by a period of depression accompanied by relative stagnation in many branches of industry. Factories are closed down, men are out of work in large numbers, and bank funds have accumulated. On all sides there are things to sell, whether goods, services or credit, and sellers are more eager than buyers; in short, at the beginning of a revival, as throughout the period of stagnation, a buyer's market obtains. There persists, too, among buyers a hesitancy to make commitments until it seems certain that both the prices of the things they buy will not go materially lower and that there will be a market for manufactured products at prices that will yield profits. When business does begin to pick up, it probably never improves at the same time or at the same rate in all industries. One group of industries may start the upward movement and for a considerable time set the pace. In the United States, for example, the textile industries normally tend to revive earlier than most other industries and did so in the present cycle; in the case of the paper industry, certain studies that have been made seem to show a tendency for at least certain parts of that industry to lag behind general business. In the present cycle, the stagnation in the iron and steel industry was almost, if not quite, unprecedented. Then almost simultaneously there developed railroad equipment purchases on an exceptionally large scale, an entirely unexpected volume of automobile production and sales, an enormous amount of building construction, and a marked demand for steel for oil tanks to store vast quantities of oil that were piling up as the result of a rate of oil production far in excess of industrial consumption. The improvement in the iron and steel industry was therefore sharp and altogether notable. Other industries have followed the lead of iron and steel, but the rate of improvement has been far from uniform. Retail trade, dealing as it does with the ultimate consumer, tends to lag behind wholesale trade, which must gamble on and therefore anticipate changes in retail business. Wholesale prices of commodities similarly tend to rise before retail prices and still further in advance of wages; and while interest rates on bonds or other long-term securities tend to decline at about the same time as interest rates on short-term bank loans, an advance in yields on bonds tends to lag behind the advance in interest rates on commercial paper, and, except for prices of real estate and other fixed capital, tends to be one of the last categories of "prices" to rise with returning prosperity. Furthermore, the various phenomena of each phase of the business cycle take place at different times in various countries and even at different times in the various sections of this country, partly as a natural effect of the regional distribution of industry.

The foregoing discussion of some of the features and in particular the irregular sequences which are found in the period of revival does not need a modified repetition to fit the other phases of the business cycle. The comments in outline form on the chart (Fig. 1) sufficiently describe the salient characteristics of each of the phases of prosperity, liquidation and depression. The essence of the period of prosperity is optimistic and feverish activity with accompanying loss of efficiency in the use of labor and capital, a seller's market with rising prices which, however, toward the end of the phase leave decreasing margins of profit; and at the peak of prosperity a definitely over-extended credit situation. The essence of the crisis or liquidation period is its short duration and the severe as well as rapid decline in business activity and in prices (at least in prices at wholesale), and a sharp increase in business failures, as those concerns which for any reason whatever are in a weak condition, find themselves unable to make it go and to meet their obligations.



I. The Cumulation of Prosperity.

1. Conditions at beginning of revival.
    - a. Considerable unemployment.
    - b. Small volume of business.
    - c. Labor and management efficient.
    - d. Low raw material, labor, and interest costs.
    - e. Debts liquidated.
    - f. Accumulated shortage of goods.
2. Favorable outlook for investment construction.
3 Volume of construction increases.
4. Volume of general business increases.
5. Prospect of rising markets stimulates forward buying.
6. Diffusion and cumulation.

II. Prosperity Contains Seeds of Depression.

1. Labor fully employed at high wages.
2. Efficiency of labor and management decreases.
3. Cost of doing business increases.
4. Selling prices increase, but not enough to maintain profit margins.
5. Stocks of goods become large and markets are overbought.
6. Investment construction falls off.
7. Tension in the money market increases.
8. Creditors begin to press for payment.

III. Liquidation and Crisis.

1. Profits decline.
2. Goods forced on market at reduced prices, buying restricted, volume of business decreases.
3. Retrenchment becomes general, unemployment grows.
4. Liquidation spreads and cumulates.
5. Prices decline more rapidly.
6. Credit strain increases, volume of business decreases.
7. Failures increase, crisis or panic follows.

IV. Depression Breeds Prosperity.

1. Credit strain is reduced.
2. The volume of business low, buying for immediate requirements only, wages fall, efficiency increases.
3. Prices and cost of doing business decline.
4. Cost of construction declines.
5. Merchandise stocks reduced, shortage of both producers' and consumers' goods gradually accumulates.
6. Credit entanglements straightened out, interest rates continue to decline.

FIGURE 1
THE FORCES OF THE BUSINESS CYCLE

The period of depression is perhaps that one about which it is most difficult to generalize. Particularly is this so in regard to the length of the period. All the difficulties of the crisis stage are likely to carry over into the depression and to persist through a good part of the latter period. The crisis comes because business tries to overreach itself; too many business men are too optimistic as to their ability to carry through indefinitely commitments on a scale which the community's consuming power cannot for a protracted period justify. Some one or more sets of causes then attack the weaker elements of the industrial structure, and as these latter meet with disaster, they drag others down with them. A serious crop failure, a sudden curtailment of buying by foreign customers, difficulties in some one large domestic industry which had been over-extended, an important strike—such events or circumstances may be responsible for the actual break of the crisis. The period of depression, however, is a time of testing for the strong and weak alike. Pessimism reigns. Few, if any, industries are likely to be exempt from the serious effects of prolonged stagnation, and the chief aim and task of a business enterprise, perhaps, must be to so completely put its house in order that at the first signs of a more favorable outlook, nothing shall have been left undone which will enable the business to take full advantage of the opportunities that offer for business expansion. Cheap and ready bank credit will be at the service of a thoroughly solvent concern to enable it to purchase raw materials when they should be bought, and long-term financing for construction purposes can be put through on a favorable basis when the proper time for expansion seems to be at hand. Labor will be

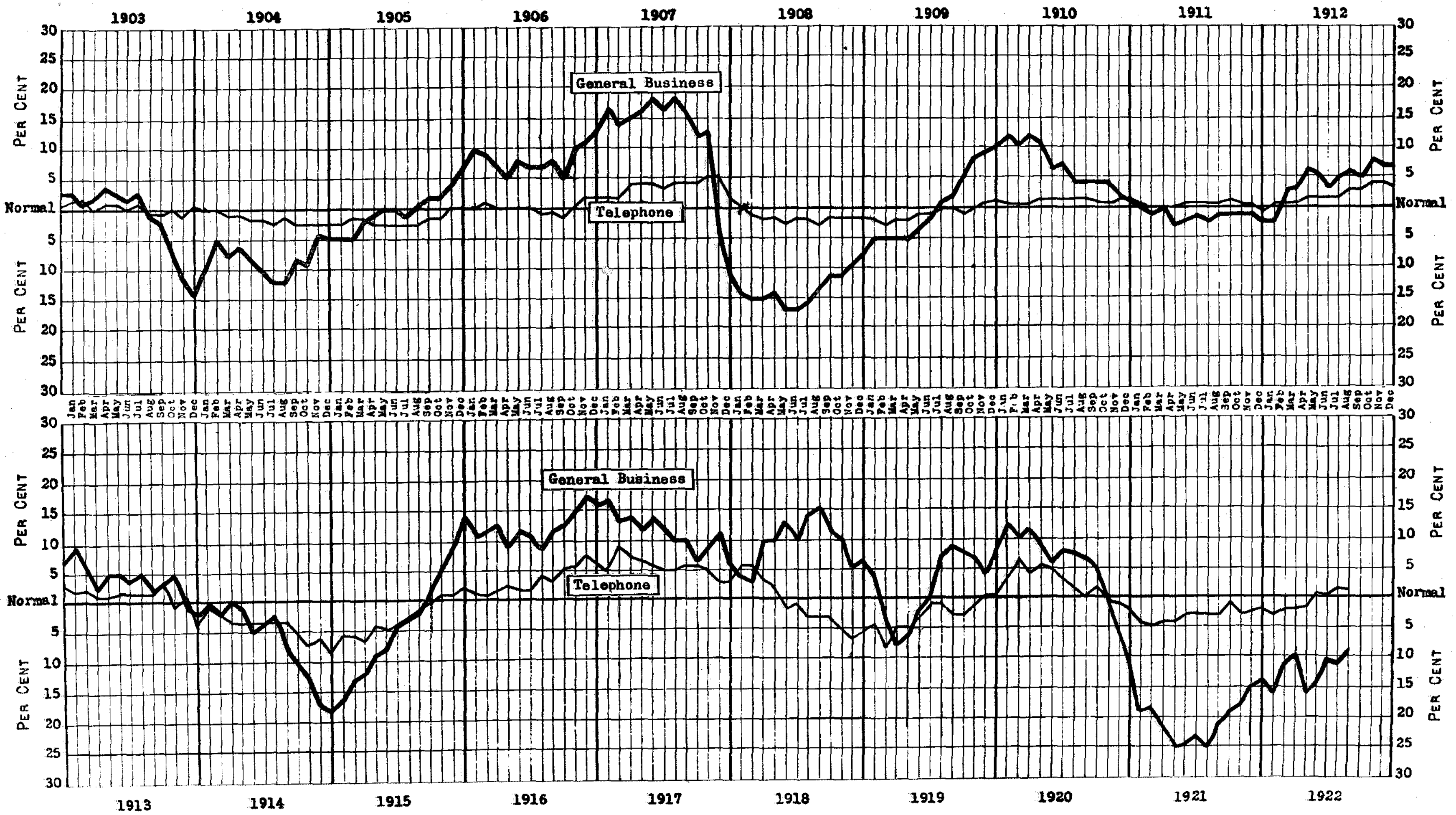


Figure 2

A Comparison of the Telephone Business and General Business in the United States with Estimated Normal

The curve "General Business" is a composite of important indices of general business activity in terms of volume of business, *i.e.*, with the effect of price fluctuations eliminated. The "Telephone" curve is a measure only of the volume of telephone business, from the standpoint of traffic, and is not significant of financial results of operation plentiful, at wages which, on the average, are below those that obtained at the peak of prosperity.

CONSTRUCTION OF A GENERAL BUSINESS CURVE

It is the all-important problem of determining what is the proper moment for increasing commitments on the eve of a revival of prosperity, or for playing safe and adopting a conservative attitude in the midst of prosperity, that has led to and stimulated the endeavor to study and interpret the various phases and features of the business cycle. To this end, curves designed to record the progress of business conditions above and below an assumed normal have been and are being constructed. The construction and use of such curves does not imply a mechanistic view of a set of such thoroughly human activities as business enterprise represents, nor do their authors consider that the graphs are a substitute for sound business judgment. It is assumed, however, that in spite of the dissimilarities between various successive business cycles and between the causes operating to affect business conditions, at different times, such graphic comparison with previous cycles as a general business curve affords, is of real value. The principal assumption underlying the construction of such a curve is that there are available certain statistics which can be used as rather faithful indices or reflectors of the scale of general business activity. The supreme importance of the iron and steel industry in supplying the wants of almost all other industries causes monthly statistics of pig iron production to be able to function generally very well as such an index. The universal use of checks as media of payment in business transactions makes monthly figures of volumes of such transactions by checks another valuable index. This is especially true if proper corrections are made for significant changes in the price level, since as prices rise or fall, a larger or smaller volume of bank credit is required to carry on the same physical volume of business. Figures for railroad freight traffic also obviously reflect the volume of trade. Statistics of production in cerain key industries other than iron and steel may be taken as further indicators of business activity. If for each of the series of figures chosen, such as that for pig iron production, the required statistics are available continuously over a period of years, it is possible by mathematical methods to determine the apparent long-time trend of each series in question; and once this trend is established and the method of allowing for seasonal variations determined on, it is possible to calculate the approximate extent to which the figure for any particular month or year is above or below the theoretical normal. A composite of the month-by-month results for each of the series enables the construction of a general business curve, which should reflect with considerable faithfulness the relation of general business activity to what may fairly be called "normal" business.

THE TELEPHONE INDUSTRY AND GENERAL BUSINESS

It is clearly possible to construct a business curve for virtually any mechanical industry for which the necessary statistics are available, in order to show the usual relationship of activity in that industry to activity in business in general, both as to time and as to relative fluctuations above and below normal. Since variations in activity in individual industries do not always or generally coincide exactly with those of industry as a whole, a careful study of normal relationships in this field is highly desirable. A business curve has been constructed for that part of the telephone industry in the United States comprised in the Bell System and appears on the chart (Fig. 2), together with a general business curve covering the last twenty years. The telephone curve, like the general business curve, concerns itself solely with physical volume of business and not at all with financial results of operations. The most obvious thing to be noted in the study of the chart is that the telephone industry is much more stable than general business, as evidenced by the fact that the curve denoting the telephone business fluctuates much less widely about the normal line than does the general business curve. The second observation is that, despite the differences in amplitude of fluctuations, the same swings above and below normal appear for general business and for the telephone business. The third feature is that, generally speaking, there is a slight lag in the curve representing the telephone business as compared with the general business curve. This lag is not uniform and indeed is not always present; and in fact, within the telephone industry itself, certain types of traffic probably vary in volume with or even slightly ahead of changes in general industrial conditions, though the volume of traffic as a whole may lag behind general business. In 1921, the recovery in the telephone business, months ahead of that in general business, is doubtless attributable to a backed-up demand for new telephone service which arose during the period of wartime restrictions upon the extension of telephone facilities and which by 1921 was being satisfied to such an extent that the cyclical turn upward in telephone traffic may be placed as early as March, 1921, while that in general business came months later.

We may say, then, that the fluctuations of general business have a definite effect on telephone business, and that the telephone business cycle is quite similar to the general business cycle, at least as to wave lengths, if not as to wave depths or amplitudes. The degree of prosperity or of adversity apparent in general business, however, for various reasons seldom gives an exact idea of conditions in the telephone business. Yet it may be said that for few, if any, industries is the study of general business conditions and of specific phases of the economic and financial situation more profitable than for the telephone business. This is true not merely, or perhaps even not principally, because an intelligent interpretation and forecast of general business conditions enables telephone executives to interpret current operating results and to plan more intelligently in regard to operating, commercial and financial problems during the six months or twelve months just ahead, but also because of the assistance it renders in planning a number of years ahead. As a public utility, the telephone industry must prepare itself to meet the demands for constantly increasing service that are put upon it. This involves planning far in advance for such additional construction, facilities and personnel as will enable the industry to meet promptly and efficiently the demands for greater service as they appear. To this end there is necessary not only an intelligent appraisal of the potential growth in telephone development, but also as definite ideas as possible of the most economical times at which to undertake construction, finance necessary expansion, and in general assume future commitments of all kinds. In all this, a careful study of present and probable business conditions, while by no means affording an infallible guide, has nevertheless been found to render valuable assistance.

Power Losses in Insulating Materials

By E. T. HOCH

Engineering Department, Western Electric Company

SYNOPSIS: *It is shown that a satisfactory measure of power loss in a dielectric is the product of phase angle and dielectric constant. Although the dielectric constant need not be explicitly considered in the design of condensers, it is important in such cases as the design of apparatus panels, and vacuum tube bases. The method used in measuring phase angle and dielectric constant is discussed.*—EDITOR.

IN working with electrical circuits operating at very high frequencies and moderately high voltages, such as radio transmitting circuits, it is found that failure in the insulation is seldom due to puncture or flashover as is usually the case at power frequencies, but is generally due to excessive heating which, in turn, causes both mechanical and chemical disintegration. As this heating is due almost entirely to the energy losses occurring in the dielectric itself, it is essential that the factors involved in the calculation of these losses be well understood.

In the past, various indices have been used as a measure of power losses for the purpose of comparing different dielectrics. Of these, power-factor, phase difference and watts per cubic centimeter probably are the most common. None of these, however, is very satisfactory for this purpose since the first two give only part of the desired information, and the last is not in any sense a property of the material, as it is dependent on both the voltage gradient and the frequency.

However, it can be shown that the product of the phase difference and the dielectric constant of a material is to a sufficient approximation an index of its power losses. Let us consider for a moment the complete expression for dielectric loss. In any condenser the capacity

$$C = a\,K$$

where a is a constant depending on the geometrical dimensions, and K is the dielectric constant. If a voltage, E, is applied to the condenser the power loss

$$P = E\,I\,\sin\Psi,$$

where I is the current through the condenser and Ψ is the phase difference of the dielectric: $\sin\Psi$ being the power factor. (Plate resistance assumed negligible.) For small angles this may be written

$$P = E\,I\,\Psi,^1$$
$$= 2\pi f\,E^2\,a\,K\,\Psi,$$

since $I = 2\,\pi\,f\,E\,C$, f being the frequency.

In the particular case of a condenser of two parallel plates

$$a = m\,\frac{A}{d},$$

where m is a constant depending on the units used, A the area of one plate, and d the thickness of the dielectric.

Hence $$P = 2\pi f\,E^2\,m\,\frac{A}{d}\,K\,\Psi.$$

But the volume of dielectric $V = A\,d$, and the voltage gradient $E_g = \frac{E}{d}$. Therefore the power loss per unit volume is

$$\frac{P}{V} = m'\,E_g^2\,f\,K\,\Psi, \qquad (1)$$

where $m' = 2\,\pi\,m$, and $m'\,K\,\Psi$ = loss per unit volume at unit frequency and potential gradient.

Thus it is seen that while no single factor of the expression can be used to represent the losses, the product of phase difference and dielectric constant[2] can be used in this way. Furthermore, for most good insulators, this product remains fairly constant throughout a considerable range of voltage and frequency. For example, we have found that for such materials as wood, phenol fibre, and hard rubber, the change of this product with frequency is of the order of 20 per cent from 200,000 cycles to 1,000,000 cycles. Hence it is possible to compare directly the losses in different materials even though the measurements were not made at exactly the same frequency.

---

[1] The substitution of the angle for its sine is correct to better than 5 per cent for angles as large as 30°.

[2] This relation has been brought to the attention of the Committee on Electrical Insulating Materials of the American Society for Testing Materials and is included in their "Tentative Method of Test for Phase Difference (Power Factor) and Dielectric Constant of Molded Electrical Insulating Materials at Radio Frequencies.

If Ψ is taken in degrees, E_g in volts per centimeter, and f in cycles per second, the constant m' reduces to $0.97 \times 10.^{-14}$. Hence, for a frequency of 1,000,000 cycles per second and a potential gradient of 10,000 volts per centimeter, the product of K and Ψ (in degrees) is within 3 per cent of being numerically equal to the dielectric loss in watts per cubic centimeter.

Data showing the variations with frequency and temperature of the phase difference, dielectric constant, and their product, for several materials are given in Tables I. and II. below.

TABLE I.

Dielectric Constant, Phase Difference and Their Product for Several Commercial Insulating Materials[3]

Material	Frequency C. P. S.	Dielectric Constant	Phase Difference Degrees	Product
Phenol Fibre A	295,000	5.9	2.9	17.1
	500,000	5.8	2.9	16.8
	670,000	5.7	2.9	16.5
	1,040,000	5.6	3.3	18.5
Phenol Fibre B	190,000	5.8	2.2	12.7
	500,000	5.6	2.5	14.0
	675,000	5.6	2.6	14.6
	975,000	5.6	2.8	15.7
Phenol Fibre C	200,000	5.4	2.1	11.3
	395,000	5.4	2.2	11.8
	685,000	5.3	2.3	12.2
	975,000	5.2	2.4	12.5
Phenol Fibre D	194,000	5.4	4.2	22.7
	500,000	5.2	3.9	20.3
	695,000	5.2	3.9	20.3
	1,000,000	5.1	3.8	19.4
Wood (Oak)	300,000	3.2	2.1	6.7
	425,000	3.3	2.0	6.6
	635,000	3.3	2.2	7.3
	1,060,000	3.3	2.4	7.9
Wood (Maple)	500,000	4.4	1.9	8.4
Wood (Birch)	500,000	5.2	3.7	19.2
Hard Rubber	210,000	3.0	.5	1.5
	440,000	3.0	.5	1.5
	710,000	3.0	.5	1.5
	1,126,000	3.0	.6	1.8
Flint Glass	500,000	7.0	.24	1.68
	720,000	7.0	.24	1.68
	890,000	7.0	.23	1.61
Plate Glass	500,000	6.8	.4	2.7
Cobalt Glass	500,000	7.3	.4	2.9
Pyrex Glass	500,000	4.9	.24	1.18

TABLE II.

Variation with Temperature of Dielectric Constant, Phase Difference and Their Product for Some Commercial Insulating Materials (Frequency 500,000 C. P. S.)[4]

Material	Temperature Degrees C	Dielectric Constant	Phase Difference Degrees	Product
Molded Phenol Product A	21	5.6	3.1	17.4
	71	6.9	6.5	45.0
	120	10.4	22.0	230.0
	21	5.5	2.9	16.0
Molded Phenol Product B	21	5.2	2.3	12.0
	71	6.1	3.7	22.5
	120	7.6	8.9	68.0
	21	5.2	2.3	12.0
Molded Phenol Product C	21	5.3	2.8	14.8
	71	6.1	3.6	22.0
	120	6.7	9.6	64.1
	21	5.0	2.5	12.5
Phenol Fibre B	21	5.6	2.5	14.0
	71	6.6	3.1	20.5
	120	6.5	4.6	30.0
	21	5.4	2.4	13.0
Phenol Fibre C	21	5.4	2.3	12.4
	71	6.0	3.9	23.5
	120	5.3	4.9	26.5
	21	4.9	2.4	11.8
Phenol Fibre D	21	5.2	3.9	20.3
	71	6.6	6.9	46.0
	120	6.3	13.5	85.5
	21	5.1	3.1	15.8
Hard Rubber	21	3.	.5	1.5
	71	3.1	1.2	3.7
	120	3.2	3.7	11.8
Pyrex Glass	20	4.9	.24	1.18
	74	5.0	.4	2.0
	125	5.0	.7	3.5
	19	4.9	.25	1.22

The above data were obtained by the resistance variation method,[5] Fig. 1. Each value of phase difference and of dielectric constant represents the average of at least five readings

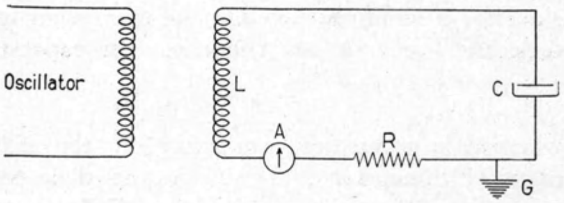


FIGURE 1

---

[3] All of the samples had been in the laboratory for some time during summer weather without artificial drying or other special preparation.

[4] The measurements on each sample were made in the order in which they are given in the table.

[5] Bureau of Standard Circular No. 74, p. 180.

on a single sample using not less than three different values of the known resistance R. A condenser, the dielectric of which consists of the material to be tested, is connected in series with a suitable inductance, a known resistance which can be varied, and a radio frequency ammeter. An oscillator is coupled loosely to the inductance and its frequency varied until resonance is obtained as indicated by maximum current through the meter. Without changing the tuning, the resistance is changed and a second reading of current is obtained. Then, since the e.m.f. induced in the measuring circuit is the same in both cases, if R_1 and R_2 are the known resistances, I_1 and I_2 the corresponding currents, and r the resistance of the remainder of the circuit,

$$I_1 (r + R_1) = I_2 (r + R_2),$$

$$r = \frac{R_2 I_2 - R_1 I_1}{I_1 - I_2},$$

or, if R_1 be made zero

$$r = \frac{R_2}{\frac{I_1}{I_2} - 1}.$$

A standardized variable air condenser having negligible resistance is then substituted for the condenser under test and the process repeated except that the circuit is tuned to resonance by varying the capacity instead of the frequency. In this way the resistance of the circuit exclusive of the test condenser may be determined. The difference between these two circuit resistances is the resistance of the test condenser from which the phase difference may be computed. The capacity of the test condenser is equal to the capacity of the standard condenser which produces resonance. From this and the dimensions of the sample, its dielectric constant may be computed.

In addition to the general precautions mentioned in the Bureau of Standards Bulletin, two others should be observed in the measurement of dielectrics. First, the electrodes must be in intimate contact at all points with the surface of the sample, as a very small air space will cause a large error in the values of phase difference and dielectric constant obtained for the sample. For this reason only mercury electrodes have been found suitable, the sample being floated on a pool of mercury forming the lower electrode and the upper electrode being formed by pouring a pool of mercury inside a metal ring on the upper surface of the sample. This introduces the second difficulty. The lower electrode being, of necessity, larger than the upper one, the electric field spreads out considerably beyond the edges of the upper electrode and increases the effective area by an unknown amount.

To determine the magnitude of this error, measurements were made on samples prepared as shown in Fig. 2. A is the upper electrode, B is the sample under test and C is a tinfoil guard ring shellaced to the lower surface of the sample but separated from the lower electrode D by a sheet of paper shellaced over the guard ring. The guard ring covers all of the lower surface of the sample except the area equal to the upper electrode and directly under it. The direct capacity between A and D is measured using the guard ring as a shield to intercept the flux which spreads out around the upper electrode, diverting it away from the measuring circuit. This measurement is made at audio frequency on a completely shielded substitution bridge. The difference between this capacity and the capacity of A to D without the guard ring is approximately the correction due to this edge effect. Using samples 6 inches square with an upper electrode 5 inches square, this correction was found to be about 7 per cent for samples ⅛ inch thick and 14 per cent for samples ¼ inch thick. All values of dielectric constant given above have been corrected. The phase difference is not affected appreciably since it is dependent only on the ratio of resistance to reactance and does not involve the area of the sample.

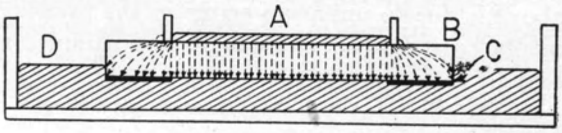


FIGURE 2

The radio frequency generator used consists of a vacuum tube oscillator having a maximum output of 250 watts. The coupling between the generator and the measuring circuit was very loose and care was taken to avoid capacity couplings. The measuring circuit was shielded from the observer by a metal screen. In spite of these precautions the results obtained on the same sample at different times do not agree as well as might be desired although the individual readings taken at the same time agree in most cases to within 5 per cent. Other observers have found that measurements made on the same sample at intervals of a few hours often differ by more than the apparent error of the measurements and have attributed it to actual changes in the properties of the material.[6] Hence it is possible that at least part of the apparent variation with frequency shown above is due to unknown errors in the measurements or unknown changes in the samples or both.

As an illustration of the error involved in taking only phase difference as a measure of power loss, suppose we wish to compare hard rubber having a phase difference of about 0.5 degree and a dielectric constant of 3., with a certain grade of glass having a phase difference of about 0.3 degree and a dielectric constant of 7. On the basis of phase difference alone the hard rubber appears very much worse than the glass, but when the dielectric constant is taken into account, the glass is found to give 40 per cent higher power loss. Similarly, some untreated woods were found to have considerably lower losses than the phenol fibres although their phase differences are nearly the same.

CONCLUSION

In the case of ordinary insulation where the object is to provide a mechanical separator or support, the product of phase difference and dielectric constant is a true measure of the energy loss per unit volume as shown by the equation (1). In the case of a *condenser* where the object is to obtain a given *capacity*, the phase difference alone determines the power loss since in this case the effect of the increased dielectric constant is exactly balanced by the smaller volume of dielectric required.

---

[6] R. Mesing—*L'Onde Electrique*, April, 1922, p. 235 and Augustin Frigon—*Comptes Rendus*, May 22, 1922, p. 1339.

The Heaviside Operational Calculus

By JOHN R. CARSON
Engineering Department, Western Electric Company

SYNOPSIS: *The art of electrical communication owes a great and increasingly recognized debt to Oliver Heaviside for his work in the developing and emphasizing a correct theory of electrical transmission along wires and in particular for his insistance on the importance of inductance. His operational methods of solving the differential equations which are fundamental of the theory of electric circuits, although not widely known, are important. These methods are peculiarly applicable to many important problems of electrical transmission. The present paper, while theoretical in character, therefore deals with a subject of practical importance to the communication engineer.*

Without attempting to give any adequate idea of the striking originality and ingenuity of Heaviside's methods, his operational calculus may be very briefly explained as follows. Problems in electric circuit theory are described by a set of differential equations involving the differential operator $\frac{d}{dt}$. *These differential equations may be reduced formally to algebraic equations by replacing the differential operator by the symbol* p *and by this expedient a purely symbolic solution is obtained. This symbolic solution is called the operational formula of the problem.*

In order to interpret the purely symbolic operational formula, Heaviside proceeded as follows: By direct comparison of the operational formula of specific problems with their known explicit solutions he was led to assign a definite significance to the operator p. *Thereupon, he obtained by induction generalized specific criteria or rules for solving the operational formula.*

The present paper, by attacking the problem from a different standpoint, shows that the Heaviside operational formula is a shorthand equivalent of an integral equation from which the methods and rules of his operational calculus are deducible.—EDITOR.

A VERY interesting and by no means the least valuable part of Heaviside's researches relates to operational methods of solving the differential equations of a class of physical problems of which electric circuit theory problems are typical; in fact Volume II of his Electromagnetic Theory is almost entirely devoted to this subject. The methods of solution which he originated and employed are of extraordinary directness and simplicity in a very large class of problems in applied mathematics. In fact it would be difficult to exaggerate the value of his work along this line, and nowhere is it more immediately and usefully applicable than in the theoretical problems of electro-technics.

Heaviside is, however, by no means easy reading and, in spite of the considerable number of published studies relating to his operational calculus, it is less generally understood and applied than its value warrants. The writer has had occasion to apply Heaviside's methods quite extensively in electrical problems and in the course of his study was led to a general formula which to him at least has proved useful in interpreting and rationalizing the operational calculus. This formula is presented and discussed in the present paper with the hope that it may be of service to students of Heaviside in understanding and applying his methods. The paper may also serve as a brief recapitulation of some of the outstanding methods applicable to the solution of problems in electric circuit theory.

The problems to which Heaviside applied his operational calculus relate to the oscillations of electrical and mechanical systems which can be described by a system of linear differential equations with constant coefficients or to such partial differential equations as *the telegraph equation*. The system is supposed to be set into oscillation *from a state of equilibrium* by suddenly applied forces and the operational formula gives the resulting behavior of the system.

The type of problem to which the operational calculus is applicable and Heaviside's method of solution may be illustrated in a sufficiently general manner for didactic purposes in connection with the solution of the system of equations

$$\begin{array}{c} a_{11}x_1 + \ldots + a_{1n}x_n = F_1(t) \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ a_{n1}x_1 + \ldots + a_{nn}x_n = F_n(t) \end{array} \qquad (1)$$

where the coefficients ajk are in general polynominals of the form

$$a_{jk} = \alpha_{jk} + \beta_{jk}\frac{d}{dt} + \gamma_{jk}\frac{d^2}{dt^2} + \ldots$$

and the α, β, γ coefficients are constants.

We are concerned with the determination of the variables $x_1 \,..\, x_n$ as functions of the independent variable t for the following boundary conditions; the known functions $F_1 \,..\, F_n$ *and the variables* $x_1 \,..\, x_n$ *are identically zero for values of the independent variable* $t \angle 0$. In other words, the system is initially in a state of equilibrium when the "forces" $F_1 \ldots F$, are applied. These boundary conditions are extremely important in physical problems.

Owing to the linear character of the equations we may without loss of generality set all the F functions equal to zero except one, say $F_1(t)$, and write

$$\begin{array}{l} a_{11}x_1 + \ldots + a_{1n}x_n = F(t) \\ \cdots\cdots\cdots\cdots\cdots\cdots\cdots \\ a_{n1}x_1 + \ldots + a_{nn}x_n = 0. \end{array} \tag{3}$$

The solution of equations (3) for the prescribed conditions may be made to depend on the auxiliary equations in the auxiliary variables $h_1 \ldots h_n$.

$$\begin{array}{l} a_{11}h_1 + \ldots + a_{1n}h_n = 1 \\ \cdots\cdots\cdots\cdots\cdots\cdots\cdots \quad (t \geq 0) \\ a_{n1}h_1 + \ldots + a_{nn}h_n = 0. \end{array} \tag{4}$$

The function on the right hand side, written, in accordance with the Heaviside notation, as unity is identically zero for $t<0$ and unity for $t \geq 0$ and $h_1 \ldots h_n$ are identically zero for $t<0$.

It may then be shown that[1]

$$x_j = \frac{d}{dt}\int_0^t F(t-y)\, h_j(y)\, dy, \quad (j=1, 2, \ldots n) \tag{5}$$

so that the solution of (3) depends entirely on (4).

Equations (4) formulate the problem actually dealt with by Heaviside who did not explicitly consider the more general equations. His method of attack was as follows; Writing p^n for the differential operator $d^n/\, dt^1$ equations (4) become formally algebraic and yield a purely symbolic solution

$$h_j = \frac{1}{H_j(p)}. \tag{6}$$

Equation (6) is the Heaviside operational formula; as it stands, however, it is purely symbolic and the problem remains to find the significance of the equation and to deduce therefrom the value of $h = h(t)$ as a function of t.

Heaviside's method from this point on was one of pure induction. From the known solution of specific problems he inferred general rules for expanding and interpreting the operational formula: the body of rules thus developed for solving the operational equation may be appropriately termed the Heaviside Operational Calculus.

The contribution of the present paper to the theory of the Heaviside operational calculus depends on the following proposition and its immediate corollary.[2]

The differential equations (4), *subject to the prescribed boundary conditions, may be written as:*

$$h_j = 0, \qquad \text{for } t<0 \text{ and } j = 1, \ldots n,$$

$$\frac{1}{pH_j(p)} = \int_0^\infty e^{-pt} h_j(t)\, dt, \qquad \text{for } t \geq 0. \tag{7}$$

The integral equation is an identity for all positive real values of p and consequently determines $h_j(t)$ uniquely.

It follows as an immediate corollary that the Heaviside operational equation

$$h = 1/H(p) \tag{8}$$

is merely a shorthand or symbolic equivalent of the integral equation

$$\frac{1}{pH(p)} = \int_0^\infty e^{-pt} h(t)\, dt. \tag{9}$$

The significance of the operational equation and the rules of the Heaviside operational calculus are therefore deducible from the latter equation. The whole problem is thus reduced to the purely mathematical problem of solving the integral equation.

It should be remarked in passing that, while the Heaviside operational calculus has been elucidated in connection with the solution of a set of differential equations involving a finite number of variables, it is not so limited in its applications. It is applicable also when the number of variables is infinite and to such partial differential equations as *the telegraph equation*. The foregoing theorem applies also to all such physical problems where an operational formula $h = 1/H(p)$ is derivable.

Before discussing the solution of the integral equation (9) and deducing therefrom some of the rules of the operational calculus, a simple but interesting and instructive example of the way the operational formula is set up will be given.

Consider a transmission line of infinite length along the positive x axis and let it have a distributed inductance L and capacity C per unit length. Let a unit voltage be applied to the line at the origin $x = 0$ at time $t = 0$; required

---

[1] This formula has been established in previous papers. It is briefly discussed in Appendix I.

[2] See Appendix I.

the line current I and voltage V at any point x at any subsequent time t.

The differential equations of the problems are

$$L\frac{\partial}{\partial t}I = -\frac{\partial}{\partial x}V,$$

$$C\frac{\partial}{\partial t}V = -\frac{\partial}{\partial x}I.$$

Replacing $\frac{\partial}{\partial t}$ by p, we get

$$I = \sqrt{\frac{C}{L}}\ e^{-\frac{px}{v}}\ V_0,$$

$$V = e^{-\frac{px}{v}}\ V_0,$$

where $v = 1\ \sqrt{L\ C}$ and V_0 is the line voltage at $x = 0$.

Now by the conditions of the problem V_0 is zero before, unity after time $t = 0$; hence the foregoing equations are operational formulas and by (9)

$$\frac{1}{p}\sqrt{\frac{C}{L}}\ e^{-\frac{px}{v}} = \int_0^\infty e^{-pt} I_x(t)\,dt,$$

$$\frac{1}{p}\ e^{-\frac{px}{v}} = \int_0^\infty e^{-pt} V_x(t)\,dt.$$

The solutions of these equations are obviously

$$\begin{aligned} I_x &= 0 && \text{for } t < x/v, \\ &= \sqrt{\frac{C}{L}} && \text{for } t \geq x/v, \\ V_x &= 0 && \text{for } t < x/v, \\ &= 1 && \text{for } t \geq x/v, \end{aligned}$$

which are, of course, the well known solutions of the problem. The directness and simplicity of the solution from the definite integrals is, however, noteworthy.

By virtue of the foregoing analysis the Heaviside operational calculus becomes identical with the methods and rules for the solution of integral equations of the type

$$1/pH(p) = \int_0^\infty e^{-pt} h(t)\,dt \qquad (9)$$

to which brief consideration will now be given.

An integral equation is, of course, one in which the unknown function appears under the sign of integration; the process of determining the unknown function is the solution of the equation. Integral equations of the form of (9) were first employed by Laplace and may be referred to as equations of the Laplace type. More recently they have become of importance in the modern theories of divergent series and summability. The solution of a large number of integral equations of the Laplace type has been worked out; however the procedure is usually peculiar to the particular problem in hand. In this connection it is noteworthy that, from a purely mathematical standpoint, Heaviside's operational calculus is a valuable contribution to the systematic solution of this type of integral equations. That is to say, methods which he developed for the solution of his operational equation suggest systematic procedure in the solution of the integral equation (9), as might be expected from the relationship pointed out in the present paper.

As stated above a large number of infinite integrals of the type appearing in equation (9) have been worked out. Consequently the solution of (9) can frequently be written down by inspection. When this is not the case, however, the appropriate procedure is usually to expand the function $1/pH(p)$ in such a form that the individual terms are recognizable as identical with infinite integrals of the required type.

An interesting expansion of this kind and one which is applicable to a large number of physical problems is as follows:

Expand $1/pH(p)$ asymptotically in the form of the divergent series

$$1/pH(p) \backsim \sum a_n/p^{n+1}.$$

This expansion is purely formal and the series is divergent. It is summable, however, in the sense that it may be identified with its generating function $1/pH(p)$. It is also summable in accordance with Borel's definition of the sum of a divergent series by the Borel integral[3]

$$\int_0^\infty dt\ e^{-pt} \sum a_n t^n/n!$$

This suggests that these two series are equal and consequently that

$$1/pH(p) = \int_0^\infty dt\ e^{-pt} \sum a_n t^n/n!$$

---

[3] See Bromwich, Theory of Infinite Series, pp. 257–259

The solution is therefore

$$h(t) = \sum a_n t^n/n!$$

provided this series, which is called by Borel the associated function of the divergent expansion, is itself convergent. This is the case in all physical problems to which this form of expansion has been applied.[3]

The foregoing will be recognized as identical with Heaviside's power series solution, obtained by the empirical rule of identifying $1/p^n$ with $t^n/n!$ in the asymptotic expansion of $1/H(p)$.

Another form of solution of very considerable practical value depends on a partial fraction expression which can be carried out in a large number of physical problems. It is

$$1/pH(p) = a + b/p + c/p^2 + \sum A_k/(p - p_k)$$

where
$$a = \left(1/pH(p)\right)_{p=\infty},$$
$$b = \left[\frac{d}{dp}\frac{p}{H(p)}\right]_{p=0},$$
$$c = \left[\frac{p}{H(p)}\right]_{p=0},$$
$$A_k = \frac{1}{p_k H'(p_k)},$$

and $p_1 \ldots p_n$ are the roots of $H(p) = 0$.

By virtue of this expansion[4] the solution is

$$h(t) = aP + b + ct + \sum \frac{e^{p_k t}}{p_k H'(p_k)},$$

where P denotes a "pulse" at the origin $t = 0$; that is,

$$\begin{aligned} P &= \infty \quad \text{at } t = 0, \\ &= 0 \quad \text{for } t > 0, \end{aligned}$$
$$\int_0^\infty P\,dt = 1.$$

In the usual case where $a = c = 0$ and $b = 1/H(0)$, this reduces to

$$h(t) = 1/H(0) + \sum \frac{e^{p_k t}}{p_k H'(p_k)},$$

which will be recognized as the celebrated Heaviside Expansion Solution.

As illustrating the flexibility of the integral identity (9), another form of solution will be given which is often of value in practical problems where an explicit solution cannot be obtained. Suppose that $1/pH(p)$ can be written as

$$\frac{1}{pH(p)} = \frac{1}{pH_1(p)} \cdot \frac{1}{pH_2(p)}$$

and that functions $h_1(t)$ and $h_2(t)$ can be found which satisfy the equations

$$\frac{1}{pH_1(p)} = \int_0^\infty e^{-pt} h_1(t)\,dt$$
$$\frac{1}{pH_2(p)} = \int_0^\infty e^{-pt} h_2(t)\,dt$$

Then the required function $h(t)$ is given by

$$h(t) = \int_0^t h_1(t - y)\, h_2(y)\, dy \qquad \textbf{(10)}$$

by Borel's Theorem (Bromwich, Theory of Infinite Series, p. 280).[5]

As a final example of the foregoing discussion we shall consider a specific problem of some practical interest in itself and which involves Heaviside's so-called "fractional differentiation" and his resulting asymptotic solutions. The physical problem is as follows: a "unit-voltage" (zero before, unity after time $t = 0$) is applied through a terminal condenser C_0 to an infinitely long cable of resistance R and capacity C per unit length. Required the Voltage V at the cable terminals.

The operational formula of this problem is easily deduced; it is

$$V = \frac{\sqrt{p/a}}{1 + \sqrt{p/a}} \qquad \text{where } 1/\sqrt{a} = C_0\sqrt{R/C}.$$

Consequently the integral equation can be written

$$\begin{aligned} \int_0^\infty e^{-pt} V(t)\,dt &= \frac{1}{p}\frac{\sqrt{p/a}}{1 + \sqrt{p/a}} \\ &= \frac{1}{p}\frac{1}{1 + \sqrt{a/p}} \end{aligned}$$

---

[3] See Appendix II.

[4] The terms $a + c/p^2$ in this expansion were suggested by Dr. O. J. Zobel and must be included in a number of important problems in electric circuit theory.

[5] This formula is quite useful; it is applied in the solution of the last example of this present paper.

Taking the last form of $1/pH(p)$, expanding asymptotically and recognizing that

$$1/p^{n+1} = \int_0^\infty e^{-pt} t^n/n!\, dt$$

$$1/p^n \sqrt{p} = \int_0^\infty e^{-pt} \frac{(2t)^n}{(2n-1)(2n-3)\ldots 1} \frac{dt}{\sqrt{\pi t}}$$

the resulting series solution can be recognized and summed as

$$V(t) = e^{at} - \sqrt{\frac{a}{\pi}}\, e^{at} \int_0^t \frac{e^{-ay}}{\sqrt{y}}\, dy$$

$$= \sqrt{\frac{a}{\pi}}\, e^{at} \int_t^\infty \frac{e^{-ay}}{\sqrt{y}}\, dy.$$

The last expansion by repeated integration by parts leads to the asymptotic series given by Heaviside. It is easy to show, also, that the series is truly asymptotic in the sense that the error is less than the last term included.

Another mode of procedure, however, suggests itself, which, by the aid of equation (10) gives the solution directly without series expansion. We have

$$\frac{1}{pH(p)} = \frac{1}{p-a} - \frac{1}{p-a}\sqrt{\frac{a}{p}},$$

$$= \int_0^\infty e^{-pt} \Big[h_1(t) - h_2(t) \Big] dt,$$

where

$$\frac{1}{p-a} = \int_0^\infty e^{-pt} h_1(t)\, dt$$

and

$$\frac{1}{p-a}\sqrt{\frac{a}{p}} = \int_0^\infty e^{-pt} h_2(t)\, dt.$$

Consequently $h_1(t) = e^{at}$, and since

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

it follows at once from (10) that

$$h_2(t) = \sqrt{\frac{a}{\pi}}\, e^{at} \int_0^t e^{-ay} \frac{dy}{\sqrt{y}}.$$

The solution $h(t) = h_1(t) + h_2(t)$ agrees with the preceding derived from the asymptotic expansion, and is considerably more direct and simple.

It is interesting to compare this solution with Heaviside's own operational solution (Electromagnetic Theory Vol. II, p. 40) which amounts to the following. The operational formula is written

$$V = \frac{p}{p-a} - \frac{1}{p-a}\sqrt{ap}.$$

The first term is discarded altogether and the second written as

$$V = \left(1 - \frac{p}{a}\right)^{-1} \sqrt{p/a}$$

$$= \left(1 + \frac{p}{a} + \left(\frac{p}{a}\right)^2 + \ldots\right) \sqrt{p/a}.$$

Identifying $\sqrt{p}$ with $1/\sqrt{\pi t}$ and p^n with d^n/dt^n the expansion becomes

$$V = \left(1 - \frac{1}{2}\left(\frac{1}{at}\right) + \left(\frac{1}{2}\right)\left(\frac{3}{2}\right)\left(\frac{1}{at}\right) - \ldots\right) \sqrt{\frac{1}{\pi a t}},$$

which agrees with the foregoing and is the actual asymptotic expansion.[6]

The foregoing discussion is sufficient, it is hoped, to show the place of the integral formula (9) in relation to the Heaviside operational calculus. It is believed to be particularly applicable in connection with a number of questions relating to divergent series and solutions which Heaviside's work has raised and which have received too little attention from mathematicians.

APPENDIX I

A proof of the integral formula

$$1/pH(p) = \int_0^\infty e^{-pt} h(t)\, dt \tag{9}$$

can be made to depend very simply on the formula

$$x(t) = \frac{d}{dt} \int_0^t F(t-y) h(y)\, dy. \tag{5}$$

This equation may be regarded as well established and can in fact be deduced in a quite general manner by synthetic arguments. It is derived and employed in papers by the writer

---

[6] The procedure by which Heaviside arrived at the foregoing asymptotic solution is not, however, always so fortunate. For example if a terminal inductance is substituted for the terminal condenser of the preceding problem, precisely the same procedure gives an incomplete result. Heaviside recognized this and added an extra term without explanation (Elm. Th. Vol. II, p. 42) but his solution appears to be doubtful in the light of some recent work by the writer in applying the formula of the present paper to the same problem.

(Trans. A. I. E. E., 1911, pp. 345–427, and *Phys. Rev.* Feb. 1921, pp. 116–134) and is deducible at once from the work of Fry (*Phys. Rev.* Aug. 1919, pp. 115–136).

On the basis of equation (5) the deduction of formula (9), in which, however, no pretense to rigor is made, proceeds as follows;

If the function $F(t)$ in equations (3) is set equal to e^{pt}, the complete solution (5) includes the particular solution[7]

$$e^{pt}/H(p)$$

which involves t only through the exponential term. The complete solution must, therefore, admit of reduction to the form

$$x(t) = e^{pt}/H(p) + y(t) \qquad \text{(a)}$$

where $y(t)$ is the complementary solution.

Now equation (5) may be written, when $F(t) = e^{pt}$, as

$$x(t) = \frac{d}{dt}\, e^{pt}\int_0^t e^{-py}\, h(y)\, dy$$

$$= \frac{d}{dt}\left\{ e^{pt}\int_0^\infty e^{-py} h(y)\, dy - e^{pt}\int_t^\infty e^{-py}\, h(y)\, dy \right\}. \text{(b)}$$

Now the first term of the expression involves t only through the exponential term while the second term involves t through the lower limit of the integral which ultimately vanishes and therefore includes no term involving t only through the exponential. Consequently the first term of (b) is identifiable as the particular solution of (a) and by direct equation it follows that

$$1/pH(p) = \int_0^\infty e^{-py}\, h(y)\, dy \qquad \mathbf{(9)}$$

which is the required formula.

The most important restriction which is implicit in the foregoing is that in splitting up the definite integral of (5) we have tacitly assumed that $h(t)$ is finite for all values of t; a restriction which is necessary in order that the infinite integral shall be convergent for all positive real values of p. This condition is satisfied in all physical problems and therefore introduces no practical limitation of importance.

However, even when this restriction does not hold formula (9) may be valid and uniquely determine $h(t)$ if p is restricted to values which make the infinite integral convergent, or when the problem is such that $e^{-pt}\, h(t)$ is an exact derivative. As an example, suppose that

$$1/H(p) = \frac{p}{p - a}$$

where a is a real positive quantity. It may be otherwise shown that $h(t) = e^{at}$ and formula (9) becomes

$$\frac{1}{p - a} = \int_0^\infty e^{-(p-a)t}\, dt$$

which is valid when $p > a$.

APPENDIX II

The discussion in the text does not pretend to be a proof of the power series expansion in any strict sense. A more satisfactory discussion proceeds as follows:

We assume that $1/H(p)$ can be formally expanded in the series

$$\sum_0^\infty a_n/p^n$$

We shall here introduce a necessary restriction on the function $1/H(p)$. It must include no function which is represented asymptotically by a series all of whose terms are zero; that is a function $\phi(p)$ such that the limit, as p approaches ∞, of $p^n\, \phi(p)$ is zero for every value of n. The function e^p is such a function. (See Whittaker & Watson, p. 154.)

With this restriction understood, start with the integral (9) and integrate by parts; we get

$$\frac{1}{H(p)} = h(o) + \int_0^\infty e^{-pt} h^{(1)}(t) dt$$

where $h^{(n)}(t) = d^n/dt\, h(t)$.

Now let p approach infinity; in the limit the integral vanishes and

$$h(o) = 1/H(\infty) = a_0$$

from the asymptotic expansion.

Integrate again by parts; we get

$$p(1/H(p) - a_0) = h^{(1)}(o) + \int_0^\infty e^{-pt} h^{(2)}(t) dt.$$

---

[7] Provided $H(p) \neq o$. This restriction is of no consequence in physical problems, where the roots of $H(p)$ are in general complex with *real part negative.*

Now let p again approach infinity; in the limit the integral vanishes, and the right hand side, by virtue of the asymptotic expansion, approaches the limit a_1, whence $h^{(1)}(o) = a_1$. Proceeding in this manner, repeated integrations by parts establish the relation $h^{(n)}(0) = a_n$. But provided the series is absolutely convergent, then

$$h(t) = \sum h^{(n)}(o)t^n/n!$$

$$= \sum a_n t^n/n!$$

which establishes the formula.

The power series solution is applicable to a large class of physical problems and has been rigorously established under certain restrictions by other methods than that employed above (see papers by Bromwich, *Phil. Mag*, May 1920, p. 407; Fry, *Phys. Rev.* Aug. 1919, p. 115; and the writer, Trans. A. I. E. E. 1919, p. 345).

On the basis of the preceding and with the aid of formula (10), expansions of the type

$$1/pH(p) \backsim \frac{1}{\sqrt{p}} \sum b_n/p^{n+1} = \int_0^\infty e^{-pt}h(t)dt$$

which occur in physical problems, can be dealt with. For since

$$\sum b_n/p^{n+1} = \int_0^\infty dt\, e^{-pt} \sum b_n t^n/n!$$

and

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

it follows from (10) that

$$h(t) = \frac{1}{\sqrt{\pi}} \int_0^t \frac{dy}{\sqrt{t-y}} \sum b_n y^n/n!$$

$$= \frac{1}{\sqrt{\pi t}} \sum \frac{b_n(2t)^n}{(2n-1)(2n-3)\,.\,.\,1}.$$

International Western Electric Company

INCORPORATED

195 BROADWAY NEW YORK, U. S. A.

Affiliated Companies

Western Electric Company, Limited Branches: Birmingham, Cardiff, Glasgow, Leeds, Manchester, Newcastle-on-Tyne, Johannesburg, Singapore.	*London, W. C. 2, England*
Bell Telephone Manufacturing Company Branches: Berne, The Hague, Barcelona.	*Antwerp, Belgium*
Le Materiel Telephonique	*Paris, France*
Western Electric Italiana	*Milan, Italy*
Western Electric Norsk Aktieselskap	*Christiania, Norway*
Vereinigte Telephon und Telegraphen Fabrik	*Vienna, Austria*
United Incandescent Lamps and Electrical Company, Limited	*Budapest (Ujpest), Hungary*
Western Electric Company (Australia) Limited	*Sydney, Australia*
Nippon Denki Kabushiki Kaisha Branches: Osaka, Dalny (Manchuria), Seoul (Chosen)	*Tokyo, Japan*
Sumitomo Electric Wire & Cable Works, Ltd.	*Osaka, Japan*
China Electric Company, Limited Branches: Shanghai, Tientsin	*Peking, China*
Northern Electric Company, Limited Branches: Quebec, Halifax, Toronto, London (Ontario), Hamilton, Windsor, Winnipeg, Calgary, Regina, Vancouver	*Montreal, Canada*
Compania Western Electric Argentina	*Buenos Aires, Argentina*
International Western Electric Company	*Rio de Janeiro, Brazil*

Sales Offices and Agencies Throughout the World

To those interested in better communication the International Western Electric Company and its Affiliated Companies offer the facilities of their consulting engineering departments to aid in the solution of problems in Telephony, Telegraphy and Radio.

Printed in U. S. A.