The Bell System Technical Journal

January, 1925

Engineering Cost Studies[1]

By F. L. RHODES

Introduction

THE subject assigned to me in the "Notes Regarding the Program of the Conference" is "The Theoretical Principles of Economic Studies and Their Possible Application in Undergraduate Courses." With your permission, I shall digress somewhat from a literal consideration of this title. I shall not undertake to derive formulae, to set up equations and to obtain maxima and minima from them. The mathematics can readily be obtained from available sources. On the other hand, I shall attempt to outline the field for economic studies in engineering work, using illustrations drawn from telephone engineering practice.

What is an engineering cost study? When you or I reach a decision to purchase a certain pair of shoes, making a selection from an assortment ranging in price from (say) $5 to $15, we have performed, consciously or unconsciously, some of the reasoning of an engineering cost study. Among the factors influencing our decision will be the probable length of service life of different pairs, as well as the ability to extend this by an expenditure, to be made at some future time, for maintenance as represented by new soles and heels, which, perhaps, can be applied economically to a moderately costly pair but not so to the cheapest.

These two elements, depreciation and current maintenance, are factors entering into engineering cost studies but they are not all of the factors. Whether we have the necessary capital in hand, or are obliged to hire or otherwise raise it, the annual cost of the capital must be taken into consideration, and treatment of the matter of depreciation is incomplete without consideration of salvage value and cost of removal.

Thus, unless we pursue our investigation into details that are not ordinarily considered when buying shoes, it is evident that our

[1] Notes of a Talk given at the Bell System Educational Conference, August, 1924.

homely illustration, while serving to center our attention on certain important subjects to be taken up in this paper, falls short in respect of others that can not be neglected in engineering cost studies. Broadly speaking, engineering cost studies deal with the comparative annual costs of alternative projects. Frequently they also involve comparisons of expenditures to be made at different times in the future. They are of value to industrial executives in assisting them to arrive at decisions where several courses of action are open, but they are not the sole guides in arriving at decisions. No hard and fast formulae can take the place of judgment based on experience. Formulae of this nature are properly used as guides to assist judgment.

The necessity for guidance from studies of this kind arises most frequently in a growing plant. The telephone plant always has been, and so far as we can anticipate, will continue to be a rapidly growing thing.

This means that whenever an addition is to be made, the question arises, how much capacity for growth is it most economical to provide for? As an illustration of this, consider with me the problem that arises when it becomes necessary to place somewhere an underground cable. Obviously it would be uneconomical to construct an underground conduit of one duct for this cable and next year or the year after to dig up the street and lay another duct for a second cable and so on in piecemeal, hand-to-mouth fashion.

On the other hand, it would not be economical to estimate the number of cables that would be required in a hundred years, even if we could foresee the needs so far ahead with any degree of certainty, and to place at the outset sufficient ducts to care for all the cables required along that route in the next century, for in that event, the carrying charges on the idle ducts would prove much more expensive, in the long run, than would additions made at infrequent intermediate times. Somewhere between one year and one hundred years is the most economical period for which to provide duct capacity in advance. The determination of this period, based on suitable construction costs, the expected rate of growth in cable requirements, and other factors is one of the useful results obtained from an engineering cost study.

Under our organization, practically all types of plant and equipment are developed by the Central Staff. These are standardized in a range of sizes sufficient to meet all the needs of the business.

The choice of standards and sizes to meet specific situations arising in the field is made by the proper officials of the associated operating companies.

If a piece of apparatus or equipment, correctly designed within itself, is installed in the wrong place, or if a wrong size is selected, loss will result.

Questions of where to place plant and what size to employ, and when to replace existing plant constantly confront the operating engineers in the field. In the telephone business every major construction project is described in what we term an "estimate" which is nothing more or less than a detailed design for the project, embodied in drawings and specifications, accompanied by a carefully prepared estimate of its cost. These estimates originate in the Plant Departments of the Associated Companies and are really the bids of the construction forces for performing the work. These estimates pass through the hands of the Chief Engineer of the Associated Company for his scrutiny and approval before they proceed to the higher officials of that company for final authorization. The Chief Engineer considers these estimates in their relation to the general plans of the Company with reference to the growth of the business and the plant. For many years the chief of the Department of which I am a member, Vice President General John J. Carty, occupied the post of Chief Engineer of the New York Telephone Company, the largest associated company of the Bell System. I have heard him say that when, while occupying that position, an estimate for some specific piece of work came before him for review, he asked himself three questions regarding it:

1. Why do it at all?
2. Why do it now?
3. Why do it this way?

Rigorous proof sufficient to answer these three questions will justify the endorsement of any engineering project, and, furthermore, each question generally involves an engineering cost study.

Fundamental Plans

Of all the engineering cost studies that are made in connection with the telephone industry, none is more far-reaching in its effect than those involved in what we term our "fundamental plans." In order to give a fair idea of the importance of the work done under our fundamental plans, it will be necessary to describe briefly what a fundamental plan is.

In completed form a fundamental plan shows what the general lay-out of the telephone plant in a city is expected to be at some definite time, usually from 15 to 20 years in the future. It shows:

(a) The number of central office districts that will be required to provide the telephone service most economically, and the boundaries of these central office districts.

(b) The number of subscribers' lines to be served by each central office.

(c) The proper location for the central office in each district to enable the service to be given most economically with regard to costs of cable plant, land, buildings and other factors.

(d) The proper streets and alleys in which to build underground conduits in order to result in a comprehensive, consistent and economical distributing system reaching every city block to be served by underground cable.

(e) The most economical number of ducts to provide in each conduit run as it is built.

These are all very definite problems that confront the executives of our Associated Companies when plant extensions are required. Our experience has shown that our fundamental plans reduce guessing to a minimum by utilizing the experience of years in studying questions of telephone growth in order to make careful forecasts on the best possible engineering basis. A few words as to how fundamental plans are made may not be out of place.

The basis of the fundamental plan is what we term a commercial survey, which is a forecast of the future community showing the probable amount, distribution and character of the population and the probable market for various classes of telephone service.

Before making this forecast, it is important to know what are the present conditions as to population and use of the telephone service. To ascertain these facts a census of the community from a telephone point of view is made. Present telephone users are classified into:

Residence Telephones.
Business Telephones in Residence Areas.
Telephones in Business Section.

In analyzing Residence telephones all families are divided among those occupying:

(a) Private Residences.
(b) Two-family Houses.

(c) Apartments.
(d) Lodging Houses.

In each class, subdivisions are made according to the rent paid as it has been found that a close relation exists between rent and the class of telephone service used. Business telephones are divided into 20 or 30 different classes. An important factor in the forecast is the future population of the city, both as a whole and by sections.

This involves, in each particular problem, not only study of the past growth of the city in question, but also careful and detailed comparisons with the growth history of other cities where conditions have been such that the experience in those places is useful in making the prediction for the city being studied.

Having arrived at forecasts, for certain future dates, as to the number of telephone users to be provided for, where they will be located, what character of service they will require, what time of day they will call, and how frequently, and where they will call, it becomes a definite, although intricate engineering problem to determine the most economical number, size and location of buildings and switchboards and the location and size of conduit runs. All of the promising combinations of future offices and districts as indicated by experience and the geographical characteristics of the city, are laid out on working maps and the annual costs are figured. The arrangement which gives the lowest equated annual costs over the period of time for which the study is made is, in general, the one which is adopted. Fundamental plans are reviewed every few years, particularly when some major plant addition, for example, the opening of a new central office, comes up for consideration. In this way we are constantly looking ahead and following a coordinated plan; but this plan is not a rigid, fixed thing. It is modified as frequently as may be necessary to meet the constantly changing requirements. In work of this kind, future expenditures must be given greater or less weight accordingly as they are required to be made in the near future or at some more distant time. This is taken into account by equating future expenditures in terms of their present worth; that is, the sum in hand, at the present time, which, at compound interest, will be just sufficient to provide for the future expenditures when they are required.

Transmission Standards and Studies

An interesting and typical annual cost problem which arises in connection with fundamental plans is that of obtaining a proper cost balance between the circuits employed for subscribers' loops and those employed in interoffice trunk lines. The larger the wire, the better will be the talk. But it will also be more expensive. The first step in solving this problem is to decide how good the transmission must be to afford satisfactory service to the telephone using public. Our present standards are a matter of growth; the accumulated results of long and extensive experience. They are live, working standards constantly being intelligently scrutinized and, when necessary, modified. A discussion of the values of the standards employed would unduly prolong this paper. Therefore, let it suffice, at this time, to state that the telephone offices in a large city, including its environs, may be divided into metropolitan offices and suburban offices; that is, the central business offices separated from the suburban residential offices. Between subscribers in different districts suitable standards of transmission are decided upon.

Before describing this study further, reference must be made to the practical necessity for the standardization of construction materials. Subscribers' loops run in length from a few hundred feet to 3, 4 or 5 miles. If we tried theoretically to make all talks exactly equal in loudness, we should have as many different sizes of wire in our cables as there are different lengths of loop. To reduce the complexity, our cable conductors are of certain standard sizes, which experience has shown are sufficiently close together to meet the needs of the business. These standard sizes, in American Wire Gauge, are Nos. 24, 22, 19, 16, 13 and 10; the three latter not being used in subscribers' loops.

Having adopted standards of transmission and standards of cable conductor sizes, our problem is to obtain the standards of transmission with the standards of cable conductors in the most economical manner.

The method of doing this, in brief, is to figure out the annual costs which would be incurred in doing it a number of different ways and to select the way that gives the lowest annual cost. In this kind of a study, which we call a "loop and trunk" study, it has been convenient to designate the subscribers' loops by their maximum circuit resistance. Adopting this form of designation, it may be assumed, first, that all of the subscribers' loops will have an average transmitting and receiving efficiency as good or better than a 350-ohm loop; as a second assumption, that they will be as good or better than a 400-ohm loop; and, as third and fourth assumptions, 450 and 500-ohm loops, respectively. In assuming, for example, a 350-ohm loop in

No. 24-gauge cable, it is, of course, necessary that all subscribers having loops longer than the amount of No. 24-gauge cable represented by this resistance shall be put in No. 22-gauge or No. 19-gauge cable as may be required.

The transmission losses, both transmitting and receiving, are then computed for the assumed loops. The transmission losses in central office apparatus are constant and known. Subtracting the losses in the offices and in the substation loops for each assumed grade of loop from the transmission standards, leaves the amount of transmission loss which can be allowed in the interoffice trunks corresponding to each limiting grade of subscriber's loop. On the basis of this allowable transmission loss in the trunks and knowing the distances between central offices, we are enabled to fix the size of conductor required in the trunks.

Knowing the grade of loops and trunks required for each of the above assumptions, we can then compute the total annual charge of giving service according to that assumption. If the assumptions have been wisely chosen it will usually work out that the first assumption, that is, a very high grade of subscriber's loop, will not be as economical as some others, due to the relatively high cost of the subscribers' loops taken as a whole. Neither will the last assumption, that is, a very low grade of subscriber's loop, be the most economical, on account of the relatively high cost of the trunks. Somewhere between, however, there will be some assumption which will show the smallest total annual charge.

To find more precisely the most economical arrangement, the various values are plotted with the assumptions as to subscribers' loops forming one set of ordinates and the total annual cost forming the other. The point on the curve representing the lowest annual cost then indicates the proper grade of subscribers' loops to employ. In the case of the longer interoffice trunks, loading is, of course, employed. In the design of toll lines and toll switching trunks generally similar cost balancing methods are employed.

In many cases, the problem can be solved by the determination of what we term "the warranted annual charge" of transmission which may be defined as the annual cost of improving the talking efficiency of the circuit in the cheapest way by a definite small amount. By means of studies of this kind, we obtain a plant closely approximating a balanced cost condition. That is, in such a plant, a dollar can be spent in improving transmission efficiency, no more effectively in one part than in another.

Other Applications of Engineering Cost Studies

From what has already been said, it should not be inferred that the sole application of engineering cost studies is in connection with the problems arising in the operating field. The question whether or not a more efficient piece of apparatus at a higher cost is warranted enters into most of our development problems. The economies of the case lie at the root of our development work in all portions of the plant.

At this point I should like to call attention to the fact that our development work covers not only what are termed "transmission" matters, but also very important problems in switchboards, outside plant and other phases of the business.

The service which we provide is a *communication* service, which involves important problems affecting the means for connecting and disconnecting the parties as well as those other important problems, to which your attention has been particularly directed, relating to the loudness and quality of the transmitted speech.

In cable design, particularly in the case of intercity cables and interoffice trunk cables, the average separation between wires in the cable affects the electrostatic capacity of the circuits and there is a definite capacity which represents the most economical degree of concentration of the wires in the cross-section of the cable. The spacing and inductance of loading coils presents another problem in balanced costs. Even in the case of wooden poles we make use of economic cost studies.

The length of life of a pole depends upon a variety of factors, the most important of which are the character of the timber; whether or not a preservative treatment is employed and, if so, the nature of the treatment; the local climatic and soil conditions and the original size of the pole.

The strength of a pole varies with the cube of the diameter of the sound wood at the weakest section. If the original size of the pole is only slightly more than the critical size at which replacement should be made, the life of the pole will be very short, as decay will reduce the size at the ground line to the critical size within a few years. On the other hand, a pole of huge size at the ground line would have a very long life before rotting sufficiently to require replacement, but the first cost of so stout a pole might readily be so great that its annual cost would exceed that of a smaller and cheaper pole. In our specifications for poles we have constantly to bear in mind that the elimination of poles containing timber defects of one kind or another means that we are adding something to the first cost of our poles and the criterion must always be whether or not the elimination of these defects will sufficiently prolong the life of the poles to warrant the increased first cost.

There have now been placed before you several examples of problems occurring in the telephone industry in the solution of which engineering cost studies may be advantageously employed, and, probably, enough has been said to make clear the importance of this form of economic analysis.

Factors Entering into Annual Costs and Their Evaluation

Let us now consider together the principal factors entering into annual cost, and how, in the course of our work, we evaluate them.

The several factors are these:

1. Cost of money.
2. Taxes.
3. Insurance.
4. Depreciation.
5. Current Maintenance.
6. Administration.
7. Operating Costs.

Cost of Money. The operating companies of the Bell System obtain the new money that they use in extensions to their plants from the sale of their capital stock and securities—bonds and notes. Such a return must be paid the investor, by the Company, as will induce a constant flow of new capital into the business. This steady influx of new capital is required because the System can not decline to expand. It is obligated to meet the increasing needs of the public it serves. Its need for new capital is a direct result of public demand for the service it renders. The rates for service which public utilities may charge are regulated by the commissions, but neither the commissions nor the utilities can fix the worth of money. Public utilities must pay the cost of money just as they must pay the cost of labor, poles and other material. No investor can be forced to invest. If the rate is below what money is worth in the general money market, he will keep out. Utility companies must bring their offerings to a general money market and submit them, in open competition, with the offerings of undertakings of every kind requiring capital. There are two ways of getting new money:

1. From investors willing to lend. These are the bond and note holders.
2. From investors willing to become partners in ownership. These are the stockholders.

Not only do stockholders expect a higher return than bond and note holders, but if the stockholders' earnings are insufficient, the bond investor will take his money to some safer market. Taking into account the ratio which must be prudently maintained between funded debt and stock, a proper figure should be obtained as representing the average annual cost of money. This figure should not be confused with the figure that represents a fair rate of return including a margin for surplus and contingencies.

Taxes. Taxes are levied by various governmental bodies, municipal, county, state and federal, on many different bases. In some specific plant problems, taxes have to be computed to meet the conditions of the case at hand but, in general, it is sufficient to employ a percentage charge for taxes based upon the average experience.

Insurance. In the case of buildings, and equipment contained in buildings, an annual cost item to cover insurance should be included.

Depreciation. Depreciation may be defined as the using up of property in service from all causes. These causes include:

(a) Wear and tear, not covered by current repairs.
(b) Obsolescence.
(c) Inadequacy.
(d) Public Requirements.
(e) Extraordinary Casualties.

All telephone property, except land, is subject to deterioration, and the continued consumption of the investment is a part of the cost of the service which must be provided for by charges against earnings. Only a small portion of the plant actually wears out in service. Instances of this are the rotting of poles and the rusting of iron wire, a relatively small amount of which is used in the plant.

On the other hand, it has been the history of the telephone business that enormous amounts of plant have been taken out of service through no defect in their physical condition but either because they had become obsolete through the development of some more economical or efficient type of equipment, or because they had become inadequate to serve the growing needs of the business.

An example of obsolescence is the replacement of antiquated methods of distribution by more modern types. Examples of inadequacy are the replacement of open wires by cable, and the replacement of small cables by larger ones. Examples of public requirement are the abandonment of pole lines and their replacement by underground construction due to road improvements, and the rebuilding of sections of underground conduit due to changes in the grade of streets or to the construction of transit subways. Examples of extraordinary casualties are fires, sleet storms and tornadoes.

The annual charge for depreciation is an amount which, if entered in operating expenses each year during the service life of a unit of plant, would, at the end of that service life, yield a sum equal to the total depreciation of that unit; that is, its first cost in place less the net salvage obtained at its removal. The consumption of capital is a necessary part of the cost of furnishing service and must be provided for by charges against earnings during the life of the property. In arriving at this depreciation charge the best thing we can do is to take our experience of years and look over the whole situation and apply our judgment to it. The value of this judgment depends on the experience, knowledge, ability and integrity of the people who exercise it.

The amount of this charge should be determined for each broad class of plant and it depends upon the average service life and the net salvage value. Net salvage value is gross salvage value minus cost of removal, and takes into consideration both value for reuse and junk value. For instance, the net salvage value of station apparatus is relatively high because a large part of the equipment can be reused in another location. In other cases, such as iron wire, the net salvage value may be a minus quantity, as there is little or nothing to offset the cost of removal.

Current Maintenance. Current maintenance charges comprise the cost of repairs, rearrangements and changes necessary to keep the plant in an efficient operating condition during its service life. In cost studies, current maintenance charges should be derived from experience and expressed, generally, on a unit of plant basis, as. for example, per pole, per mile of wire, per foot of cable, or per station, according to the kind of plant being considered. Generally speaking, they bear no direct relation to first cost of plant as other annual charges do.

For this reason, when comparing the annual costs of two or more plant units of different sizes or types, an incorrect result would be obtained if maintenance charges were expressed as a percentage of the first cost.

However, for comparative cost studies of average plant, maintained under average conditions, it is sometimes within the precision of the study to employ figures expressed as a percentage of the first cost, provided the figures were derived from the cost of maintaining average plant where average conditions were known to obtain.

Administration. In certain cost studies, a small allowance is usually made to cover that portion of the salaries and expenses of the general officials of the Company which is fairly chargeable to the administration of the plant.

Operating Costs. In certain classes of engineering cost studies, comparisons may involve the situation where one type of plant costs initially more than an alternative type, but permits savings to be made in the daily operating labor which may or may not offset the additional first cost. In such cases, to obtain a true comparison, the operating labor costs under each plan must be combined with the total annual charges which are applied to the first costs of the respective plant quantities.

Present Worths

Engineering cost studies frequently involve a balance between plant installed at the present time and plant installed at some future time. An example of this would be the comparison of a pole whose life was to be extended by attaching it to a stub after (say) 15 years, with a stouter and more expensive pole installed at present or with a pole to which preservative treatment was applied prior to its installation.

In such cases it is not sufficient to compare annual costs which are to be incurred at different times without reducing them to a basis upon which they can properly be compared. If a given amount is required to be expended at some future time, it obviously requires a smaller sum at present in hand to meet this obligation if the fixed time is far distant than if it is in the immediate future.

Let us picture ourselves at the end of the year 1924. If an annual charge of $1,000 is to be paid each year for the 5 years beginning January 1, 1925 and ending December 31, 1929, there will be required, to provide for these five $1,000 payments, the sum of $4,100, in hand, assuming that interest is compounded annually at 7 per cent. On the other hand, if these five annual payments of $1,000 each instead of beginning in 1925 were to begin ten years later, that is, if they were to run from January 1, 1935 to the end of 1939, we should require, in hand, $2,084, that is, only about half as much.

To compare, upon a fair basis, expenditures that have to be made at different times, it is customary, as has been done in the preceding example, to reduce these different expenditures to their "Present Worths," or the equivalent in equated or accumulated annual charges.

Summary

From all that has been said, it becomes evident that, whenever a specific addition is made to a growing plant, we are, to a greater or less extent, committing ourselves to a definite programme for relieving, reinforcing or replacing it at some future time in order most economically to provide for the requirements of growth.

The underlying thought, which can not be overemphasized, is so to plan the plant that, as far as practicable, it will serve for its full life, and require no wholesale changes involving the abandonment of substantial portions of the installation. While the design should be based upon the best estimates of future growth that are obtainable, it must be recognized that the most carefully designed plant layouts employing the best possible estimates of growth, may not always meet the ultimate requirements of flexibility. The chances of a comprehensive plan not fitting in with future development can, however, be reduced to a minimum by thoughtful initial planning.

Generally speaking, our distributing plant layout, once it is established, can not readily nor economically be materially changed. Consequently, if it is not sufficiently flexible in the fundamentals of its design to meet reasonable future possibilities, it may affect adversely the carrying out of proper and economical relief measures, or may require abnormally early reconstruction or replacement. It is very desirable, therefore, always to keep in mind, in any plant layout work, the progressive relief steps which are likely to be required to meet the changing conditions affecting the service requirements. Whenever plant is moved, or taken out of service, property loss is realized. Certain expenditures for these purposes represent the most economical way of conducting the business. But it is of the utmost importance that they should always be incurred along the line of maximum economy, which means that behind every plant addition must be engineering cost studies to assist in furnishing the answers to the three questions:

Why do it at all?
Why do it now?
Why do it this way?

But it must always be borne in mind that these studies do not and can not, in themselves, constitute the sole criterion for determining what should be done. They are, at the best, only an aid, guide and check to be utilized, within their limitations, in arriving at conclusions that must, in the last analysis, rest upon seasoned judgment and experience.

Nevertheless, so great do we find the importance of these engineering cost studies in our work, and so great must be their importance in the engineering of any other kind of growing plant, that the question might be raised whether, in courses of engineering instruction, a few hours at least could not advantageously be devoted to acquainting the student with the nature and importance of these economic problems.

The Limitation of the Gain of Two-Way Telephone Repeaters by Impedance Irregularities

By GEORGE CRISSON

Introduction

BECAUSE of the fact that it is a difficult and expensive matter to build and maintain the high grade circuits that are required for modern long distance telephone transmission with repeaters, many workers in this field have attempted to devise some form of two-way repeater which would be able to give as large a gain as desired without singing or poor quality due to irregularities existing in the lines. They have thought that if such a repeater could be constructed it would permit the use of lines less carefully built and, therefore, cheaper than are at present required, and that fewer repeaters would be required because larger gains could be obtained at each repeater.

As a matter of fact the irregularities in the lines have a very important effect and control, to a great extent, the repeater gains which can be used whenever a telephone circuit is arranged so as to be capable of transmitting in both directions over a single pair of wires with constant efficiency.

It is the object of this paper to explain, in a very simple way, why this is true. To do this the phenomenon of electrical reflection is first made clear. Then a two-way repeater system is introduced and the effects of reflection upon this system are explained. After mentioning several of the types of repeaters which have been used successfully, the paper concludes with an explanation of the fallacies underlying a number of schemes which have been proposed from time to time by various inventors.

Reflection in Telephone Lines

Whenever discontinuities or irregularities exist in telephone circuits, reflection of a certain part of the speech wave takes place at each irregularity. In order to appreciate why it is that irregularities in two-wire telephone circuits affect very greatly the amount of repeater gain which can be secured whenever two-way operation is desired, it is first necessary to obtain a clear picture of why it is that reflections take place at irregularities.

Fig. 1 represents an infinite ideal telephone line without repeaters. If such a line is non-loaded or continuously loaded each part of it is exactly like every other part having the same length. If the line is loaded with coils then each loading section is exactly like every other loading section.

When a telephone transmitter or other signaling device A acts upon such a line it causes a wave to travel over the line away from

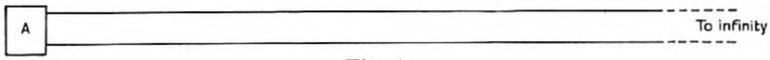


Fig. 1

the source. If the line includes resistance or other losses this wave gradually becomes smaller until it is too weak to be detected but no portion of the wave returns to the source after once leaving it.

If some portion of the line differs in its electrical makeup from other portions of the line it constitutes an irregularity and interferes with the passage of the wave.

Fig. 2 shows a line exactly like that of Fig. 1 except that an irregularity B has been introduced. This irregularity has been shown

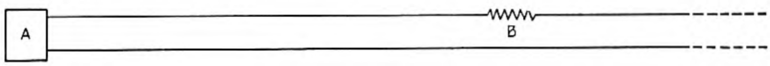


Fig. 2

as a series resistance though any other departure from the regular electrical structure of the line would produce similar effects.

When a wave encounters such an irregularity, it splits into two parts one of which continues in the original direction of propagation along the line while the other is propagated in the opposite direction toward the source.

In order to understand this phenomenon, which is called reflection, imagine that a wave is traversing the line from left to right. As it passes the point B a current flows through the series impedance which constitutes the irregularity and this causes a drop of potential through the impedance. Obviously, this changes the state of affairs as there is now a sudden alteration in the voltage across the line as the wave passes the irregularity whereas there is no such alteration without the irregularity.

Suppose that for the impedance element we substitute the output terminals of a generator which has a negligible impedance and arrange the generator so that it is excited by the wave traveling over the line but that the excitation is not affected by the voltage set up by the generator itself. Such an arrangement is shown in Fig. 3. The arrangement for exciting the generator is supposed not to require an appreciable amount of power or to constitute an irregularity. This generator then resembles the series impedance of Fig. 2 in that it produces no disturbance in the line when no waves are passing but as soon as a wave arrives the generator becomes active and produces a

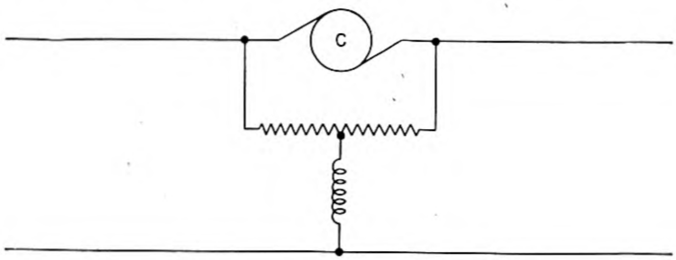


Fig. 3

voltage in series with the line. By proper adjustment of the exciting mechanism of the generator the voltage across its output terminals can be made just equal to the disturbance produced by the impedance element at B in Fig. 2 and so exactly reproduce the effects of the irregularity. In order to do this the generator might have to absorb energy from the wave passing over the line instead of giving it out, but it would establish the desired voltage relations.

Now as the generator has no appreciable impedance the wave passes through it without interference but the e.m.f. which it sets up obviously sends out waves in each direction from the generator.

On the right of the irregularity will be found one wave made up of the original undisturbed wave combined with that from the generator and traveling onward in the original direction. The combined wave will usually be smaller than the original wave though it might under some circumstances be larger and its shape might or might not be altered depending upon the nature of the irregularity and the character of the line.

On the left of the irregularity will be found the original wave traveling from left to right and the reflected wave traveling from right to left.

By a similar process of reasoning the reflection caused by bridging an impedance across the line at the point B can be illustrated. In this case the output terminals of the generator should be bridged across the line and made of very high impedance.

Any departure from the regular structure of the line such as occurs at the junction of two lines of different types or where loading coils have the wrong inductance or are wrongly spaced causes reflections in the manner described above.

Ideal Repeater on an Ideal Line

Fig. 4 shows an ideal telephone circuit consisting of two sections of line L_1 and L_2 which are free from irregularities and are joined by a repeater R. The remote ends of the line sections are connected to terminal apparatus A_1 and A_2 which have impedances which

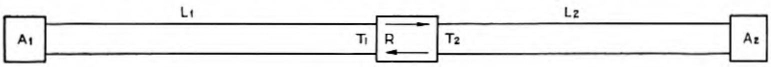


Fig. 4

smoothly terminate the lines, that is, if either line had originally extended to an infinite distance from the repeater and had been cut to connect it to the terminal apparatus, this apparatus would have the same impedance as the part of the infinite line which was cut off. The construction of the repeater R is limited only by the requirement that if an electric wave arrives at the repeater terminals T_1 or T_2 over either line a similar but larger wave is transmitted from the repeater over the other line. The gain of the repeater determines the relative sizes of the waves arriving at and departing from the repeater.

If now a wave is started at one end of the circuit, for example A_1, it traverses the line L_1 and is absorbed or dissipated in the portion of the repeater connected to the terminal T_1. This wave acts upon the internal mechanism of the repeater in such a way as to send out a larger wave which traverses the line L_2 and is completely dissipated in the terminal apparatus A_2.

Ideal Repeater on a Line Containing Irregularities

Fig. 5 illustrates a line exactly like that of Fig. 4, except that an irregularity B_1 (or B_2) has been introduced into each section. If a

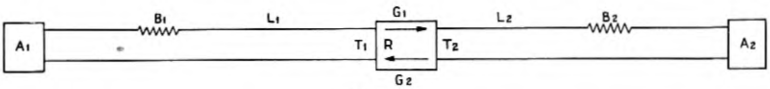


Fig. 5

wave leaves one terminal such as A_1, it traverses the line L_1 eventually arriving at the terminal T_1 of the repeater R with a certain strength. This wave is amplified and transmitted into the line L_2 which it follows until it encounters the irregularity B_2. At B_2 it is partially reflected, one portion returning to the repeater and the other traveling to the terminal A_2 where it is absorbed. The reflected wave passes through the repeater, is amplified and transverses the line L_1 until it encounters the irregularity B_1 where it is again reflected, one part being propagated to the terminal A_1 where it is dissipated, while the other part returns to the repeater and repeats the cycle of amplification and reflection. This action continues indefinitely the wave being reflected alternately from the irregularities B_1 and B_2.

If the total gain in the round trip path is greater than the total loss the wave will be stronger on each arrival at any point in the circuit than on the preceding trip and will continually increase in power until the power limits of the repeater or some other cause prevents a further increase and a steady sing is established. If the gain is less than the loss, the wave will become weaker with each trip from B_1 and B_2 and back until it falls below the strength which can be detected.

Evidently, if the repeater gain is made so great that a steady sing is established, satisfactory telephoning over the circuit will be impossible. Serious quality impairment may occur, however, when the gain is not so great as this. Consequently, when irregularities are present in a line containing repeaters, the repeater gains are necessarily limited.

In the above illustration, it was assumed that two irregularities were present. Serious effects, however, due to the production of echo effects which may be heard by the talker, may be produced by reflection from a single irregularity. Consequently, a single irregularity in the circuit will set a limitation on the repeater gain even though it could not cause singing if a 22-type repeater were used.

From the foregoing explanation, it is evident that the effect of the reflections at the irregularities, which limits the repeater gains, is not dependent upon any special properties of the telephone repeaters. These limitations will necessarily exist with any types of repeater whatsoever which have the property of producing amplification in both directions at the same time.

Effect of Using the Wrong Line Impedance

The discussion will now be extended to show that not only must the lines with which a repeater is to work be smooth, if limitation of the gains is to be avoided, but also the repeaters must be designed to fit lines of one particular type. It has just been shown that reflection takes place if a series or a bridged impedance is inserted in a line. This reflection will take place whether the impedance is inserted at some intermediate point in a line or adjacent to a repeater. Inserting such an impedance adjacent to a repeater would, on account of this reflection, seriously limit the gain which could be produced by the repeater. Now inserting an irregularity adjacent to a repeater amounts to the same thing as substituting a line having a different impedance for the line with which the repeater is designed to function. Since any change in the impedance of a line connected to a repeater away from the impedance with which the repeater is designed to work is equivalent to inserting an irregularity adjacent to the repeater, it is evident that *it is impossible to construct a repeater system whose amplification will be constant in both directions and whose gain will not be limited by irregularities in the lines and by any departure of the line impedance from that for which the repeater is designed.*

Successful Types of Repeaters

Two forms of repeater circuit, the well known 21 and 22 type circuits, have been developed to the point where they have become highly important and successful parts of the telephone plant. These have been so completely described in a paper entitled, "Telephone Repeaters" by Messrs. Gherardi and Jewett,[1] that no further description will be attempted here. It is sufficient to point out that in the case of the 22 type repeater the necessary impedance requirements are met by providing networks which imitate closely the characteristic impedances of the two associated lines. Any departure of the line impedance from the value for which the network was designed or any irregularities in the line or terminal equipment impose limits on the obtainable gain in the manner described above. In the case of the 21 type circuit the impedance requirements are met by putting the repeater between two similar lines whose impedances balance each other.

Another type of repeater circuit, called the booster circuit, was mentioned in the paper just referred to. This circuit does not depend upon impedance balance in the same way as the 21 and 22 type circuits and it is capable of giving two-way amplification but its performance is even more seriously affected by impedance deviations in the lines than the latter circuits. The booster form of repeater circuit has not yet proved useful in a commercial way.

[1] Proceedings of the American Institute of Electrical Engineers, 1919, page 1255.

Devices Employing Voice Controlled Relays

Many different devices aiming to secure the practical equivalent of two-way repeater operation by means of relays (mechanical or thermionic) controlled by the voice currents themselves have been suggested. In these devices the action of the relays is such that when transmission is passing in one direction through a repeater, the transmission in the opposite direction is either wholly or partially blocked. Evidently the gain of such a repeater as this is not limited by impedance irregularities in the lines, since it is really a one-way device during the passage of speech currents.

Repeaters controlled by voice operated devices will not be discussed here further in view of the fact that the principal object of this paper is to treat repeater systems which are truly two-way in their operation.

Other Types of Repeater That Have Been Proposed

Several of the arrangements that have been proposed by inventors who sought unsuccessfully to produce two-way repeaters not subject to limitation by line irregularities will now be described.

1. *Repeaters Involving Balance.* A great many circuits have been devised which involve the principle of balance. These always involve the same fundamental principle as the hybrid coil used in the repeaters now in commercial service though often the arrangement appears quite different. This principle is that the output energy of the amplifier working in one direction, for example, the east bound amplifier, is divided into two parts, one of which is sent into the line east and the other into the corresponding network. The input terminals of the west bound amplifier are so connected that the effect on them of the current entering the line east is opposed by the effect of the current entering the network and consequently the impedances of the line and network must accurately balance each other to keep the output energy of one amplifier out of the input circuit of the other. Sometimes the balance is effected by connecting the line and network into a common electrical circuit and connecting the input terminals of the amplifier to two points of equal potential in this circuit. In other arrangements two fluxes which depend upon the currents entering the line and network are balanced against each other in the core of a special transformer so that a winding connected to the input of the amplifier is not affected.

Usually the impedance of the network equals that of the line, but arrangements are possible and even have certain advantages in which the energy is not equally divided between the line and network and the impedance of the network is either greater than or less than that of the line in a certain ratio.

Through unfamiliarity with the principles involved the inventors sometimes assume that an approximate balance such as might be obtained by using a simple resistance is sufficient to meet all requirements. None of these arrangements, however, can avoid the effects of departures of the line impedance from the values for which the networks are designed nor can they better the performance of the present repeaters in respect to the effects of impedance departures. Usually such circuits are inferior in some important respect to the arrangements now in use.

2. *Circuits using Rectifiers.* In one type of circuit the inventors propose to use rectifiers to prevent the output energy of one amplifier

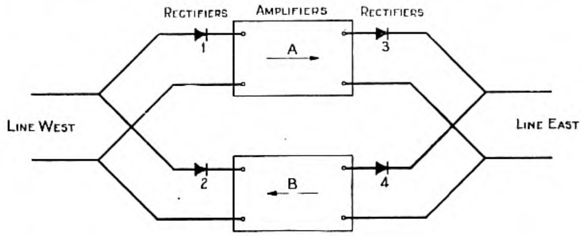


Fig. 6

acting upon the input circuit of the other. A simple diagram illustrating the operation of this scheme is given in Fig. 6. Rectifiers are placed in series with the input and output circuits of both amplifiers and poled in the directions indicated by the arrow heads which point in the direction the rectifier is supposed to permit current to pass. It is argued that the rectifier in the output circuit of each amplifier permits only currents of one polarity to enter the line and that the rectifier in the input circuit of the opposite amplifier is so poled that these output currents cannot pass it into the input circuit and, therefore, singing cannot occur.

If a wave arrives, for example over the line west, the positive half waves pass through the rectifiers 1 and 2 into the input of the east bound and the output of the west bound amplifier respectively. The negative half waves are suppressed by the rectifiers. This is illustrated by Fig. 7 which shows the wave arriving over the line and Fig. 8 which shows the part of the wave which enters the amplifiers.

That portion which reaches the output of the west bound amplifier is lost while the portion which reaches the input of the east bound amplifier, is amplified, and passed on through the rectifier 3 to the line east. If the amplifier were completely distortionless and, therefore, capable of amplifying direct currents and the rectifiers perfect, that is, offering zero resistance to currents in one direction and infinite resist-

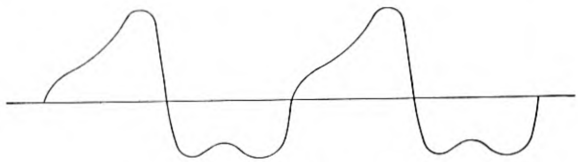


Fig. 7

Fig. 8

ance to currents in the opposite direction, the currents transmitted to the line east would have the wave shapes shown in Fig. 8.

As it would be impracticable to make the amplifier amplify the direct-current component of the wave shown in Fig. 8 the amplifier would tend to send out a wave somewhat like that shown in Fig. 9,

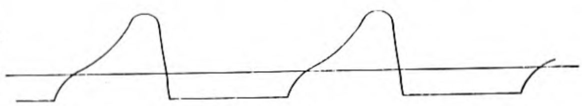


Fig. 9

which is the wave of Fig. 8 with the direct component removed. The rectifier 3 then suppresses the negative half waves, finally permitting the wave shown in Fig. 10 to pass to the line east. On account

Fig. 10

of the great distortion involved the quality of speech would be greatly impaired if, indeed, the speech would not be rendered unintelligible.

Assuming, however, that intelligible speech is possible in spite of this distortion, the rectifiers would not prevent singing. Suppose the repeater shown in Fig. 6 to be cut into the line shown in Fig. 5 at R and that waves are arriving from the line west. There are certain line conditions which are practically certain to exist and which would send back reflected waves that would reverse the potential across the line east at the terminals of the repeater, causing impulses to reach the input of the west bound amplifier. These impulses will be amplified and returned to the line west where, if similar conditions exist, they will once more enter the east bound amplifier. If the gains are great enough to offset the losses caused by the rectifiers, the system will sing.

It is, therefore, evident that rectifiers offer no chance for improving on the action of the present types of repeaters because they cause

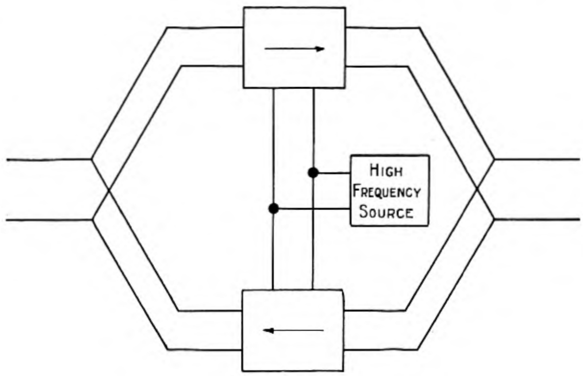


Fig. 11

serious distortion and do not prevent singing except under certain special conditions not likely to be found under practical conditions.

3. *Circuits using High-Frequency Switching.* Another device which is frequently proposed in one form or another is illustrated in Fig. 11. In this case an amplifier is provided for each direction of transmission. These amplifiers are so designed that their amplifying power can be destroyed and restored periodically at high frequency by currents from a suitable source, the amplifier in one direction being active when the other is inactive. The frequency of the controlling currents is above the audible range. In a variation of this scheme a single amplifier is used which is pointed first in one direction and then in the other at a frequency above the audible range. It is argued that since there is amplification in only one direction at any given instant the system cannot sing.

Imagine such a repeater to be inserted in the line at R in Fig. 5, and that voice waves are arriving over the line from A_1. Owing to the nature of the repeater these waves will be cut up into a series of pulses having a frequency equal to that of the controlling current and varying in magnitude according to the shape of the voice wave being transmitted. These pulses will be partially reflected at the irregularity B_2 and part of their energy will return to the repeater. Due to the fact that a finite time is required for the pulses to pass from R to B_2 and back, they are likely to arrive at the right moment to find the amplifier set for amplification in the opposite direction, in which case they will pass through towards A. For a single irregularity, it would be possible to select a frequency such that the pulse would return when the repeater is set against it, but this would require a different frequency for each irregularity which is obviously impossible.

In case the line cannot transmit the high frequency pulses, their energy would be stored in the inductance or capacity of the first elements of the line L_2 and returned to the amplifier when it is in condition to transmit from L_2 to L_1. To avoid the latter objection it has been proposed to employ low pass filters on the output side of each one-way amplifier to convert the high frequency pulses back into ordinary voice waves before passing them into the line, but this obviously defeats the object sought in using the high frequency control of the amplification because each amplifier now receives ordinary voice waves and gives out enlarged copies of them which are subject to the same reflections as if plain one-way amplifiers without the high frequency control had been used.

From these considerations it will readily be seen that repeater systems depending upon high frequency variation of the gain to avoid singing and the necessity for impedance balances are inherently unworkable.

Practises in Telephone Transmission Maintenance Work[1]

By W. H. HARDEN

Synopsis: This paper describes the practical applications of transmission maintenance methods in a telephone system. The methods applicable to toll circuits of various types are first discussed, information being included in this connection on the maintenance of the amplifier circuits involved in telephone repeaters and carrier. Testing methods applicable to the local or exchange area plant are next described, the description including both manual and machine switching systems. The results accomplished in toll and local transmission maintenance work are considered from the standpoint of the kinds of trouble which can be eliminated and the effect which these troubles have on service.

The methods described in the main body of the paper relate particularly to tests of volume efficiency. Certain other transmission maintenance testing methods directly associated with volume efficiency tests are briefly described in Appendix A of the paper.

IT is the purpose of this paper to present a general picture of the practical applications of methods of measuring transmission efficiency in the Bell System which have been developed by study and experience under plant operating conditions. The rapid growth of the telephone industry has made it necessary that these methods be such as to allow them to be applied on a large scale in a systematic and economical manner thereby providing for a quick periodic check of the efficiency of the various types of circuits as they are used in service.

Transmission maintenance can be broadly defined as that maintenance work which is directed primarily towards insuring that the talking efficiencies of the telephone circuits are those for which the circuits are designed. There are, of course, many elements which affect the talking efficiency and various d-c. and a-c. tests are available for checking the electrical characteristics of circuits and equipment to insure that these characteristics are being maintained in accordance with the proper standards. In the final analysis, however, an overall test of the transmission efficiency of the circuit in the condition it is used in service will show at once whether it is giving the loss, or in the case of amplifier circuits, the gain which it should give. Transmission tests, therefore, offer a means whereby many of the electrical characteristics of circuits can be quickly and accurately checked.

In referring to transmission testing apparatus in this paper, four standard types described in previous papers are involved. The first three types listed below were described by Best and the fourth by

[1] Paper presented at the Pacific Coast Convention, *A. T. E. E.*, October, 1924; abstracted in the *Journal*, *A. I. E. E.*, Vol. 43, p. 1124, 1924.

Clark.[2] Reference in these papers was also made to the standard oscillators used in supplying the measuring currents for the sets.

1—*A Transmission Measuring Set.* This is an "ear balance" portable set suitable for loop transmission testing only and designed primarily for testing equipment and circuits in the smaller central offices.

3—*A Transmission Measuring Set.* This is a "meter balance" portable set suitable for both loop and straightaway transmission testing and designed primarily for testing circuits and equipment in the larger central offices.

4—*A Transmission Measuring Set.* This is a "meter balance" set suitable for both loop and straightaway transmission testing and designed for permanent installation at the larger toll c[illegible]s primarily for testing toll circuits.

2—*A Gain Set.* This is a "meter balance" set designed for measuring amplifier gains.

Certain other testing methods in addition to volume efficiency tests are also extensively used in transmission maintenance work and some of the more important of these are briefly discussed in Appendix A of this paper.

Since the routine procedures in testing toll circuits using the above apparatus differ considerably from those followed in the local or exchange area plant, the toll and local practices have been considered separately in the following discussions:

Transmission Tests on Toll Circuits

The importance of having available means for quickly checking the transmission efficiency of toll circuits and of economically maintaining the proper standard of transmission is evident when it is considered that in a plant such as that operated by the Bell System there are at the present time more than 20,000 toll circuits in service. The circuits making up this system are of various types and construction, depending on the service requirements and length, and also upon certain other factors determined by engineering and economical design considerations.

From the standpoint of maintaining transmission efficiency between toll offices, the various types of toll circuits can be divided into three general classes: one, non-repeatered circuits, two, circuits equipped

[2] F. H. Best, "Measuring Methods for Maintaining the Transmission Efficiency of Telephone Circuits," *Journal of the A. I. E. E.*, February, 1924. A. B. Clark, "Telephone Transmission over Long Cable Circuits," *Journal of the A. I. E. E.*, January, 1923.

with telephone repeaters and three, circuits equipped for carrier operation. The latter two classes are alike in many respects as far as the maintenance methods are concerned and both require somewhat more attention than the circuits not equipped with amplifying apparatus. The length and number of repeaters involved are also important factors which must be taken account of in tandem repeater and carrier circuit maintenance. Very long tandem repeater circuits

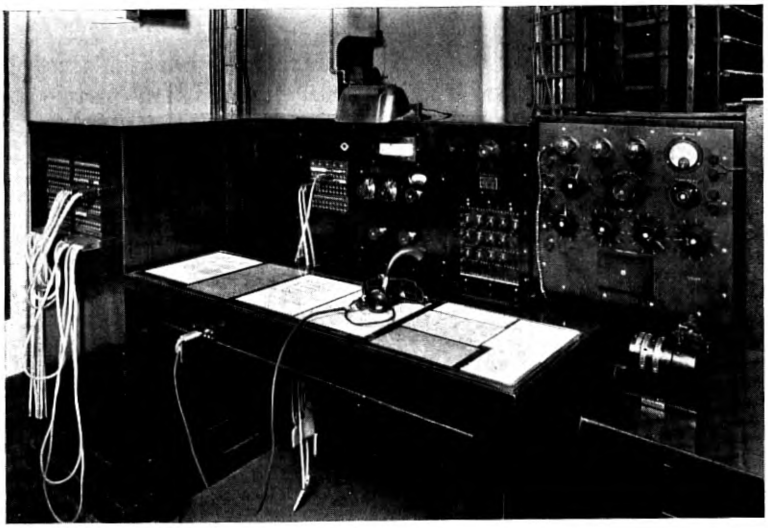


Fig. 1—Illustration of 4-A Transmission Measuring Set and 4-B Oscillator Installed in a Toll Test Room

such, for example, as the long toll cable circuits described by Clark[2] require special maintenance procedures similar in many respects to those required in carrier maintenance.

The 4-A type of transmission measuring set generally used for testing toll circuits may be considered as a toll transmission test desk. Fig. 1 shows a picture of one of the latest models together with an oscillator for supplying the measuring current, installed at a toll office for use in routine testing. The set is provided with trunks to both the toll testboard and toll switchboard, and also with call circuits to toll operators' positions for use in ordering up circuits for test. The electrical measuring circuit is designed so that tests may be made on two toll circuits looped at the distant end, or straightaway on one toll circuit the distant terminal of which terminates in an office also equipped with a transmission measuring set of the same type.

To illustrate the application of this toll transmission test desk, Fig. 2 shows schematically an arrangement of four toll offices having circuits between them of the three general classes—non-repeatered, repeatered and carrier. Offices A and D are equipped with transmission measuring sets of the type shown in Fig. 1. A logical testing

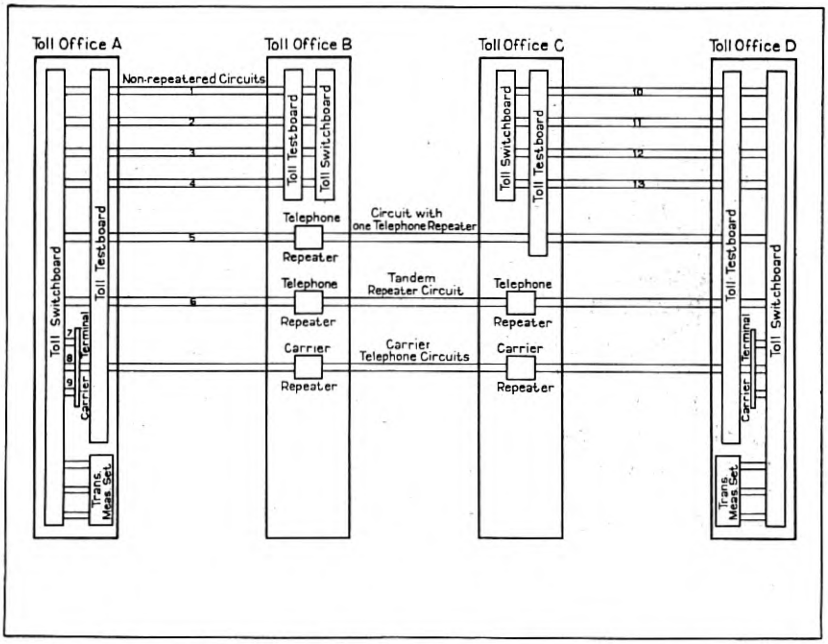


Fig. 2—Schematic Diagram of Typical Toll Circuit Layout to Illustrate General Method of Testing Non-Repeatered, Repeatered and Carrier Circuits

procedure for the arrangement in Fig. 2 is for offices A and D to test the non-repeatered circuits 1 to 4 and 10 to 13 by having them looped two at a time at the distant terminal offices B and C. By "triangulation measurements" on any three circuits in each group, the equivalent of each individual circuit can be readily computed.

For the circuits 5 to 9 extending between offices A and D equipped with telephone repeaters or carrier, straightaway measurements can be made in each direction with the two transmission measuring sets provided. Loop tests could, of course, also be made on the circuits from either office A or D, but this would require cutting the telephone repeaters out of one circuit or having available a non-repeatered or non-carrier circuit, since the gains of the repeaters in the two directions introduce variable factors in the overall equivalents which do not permit triangulation computations to be made. The overall tests on the carrier circuits do not differ in any way from the tests on repeatered or non-repeatered circuits, each carrier channel being tested as a separate circuit through the switchboards. The

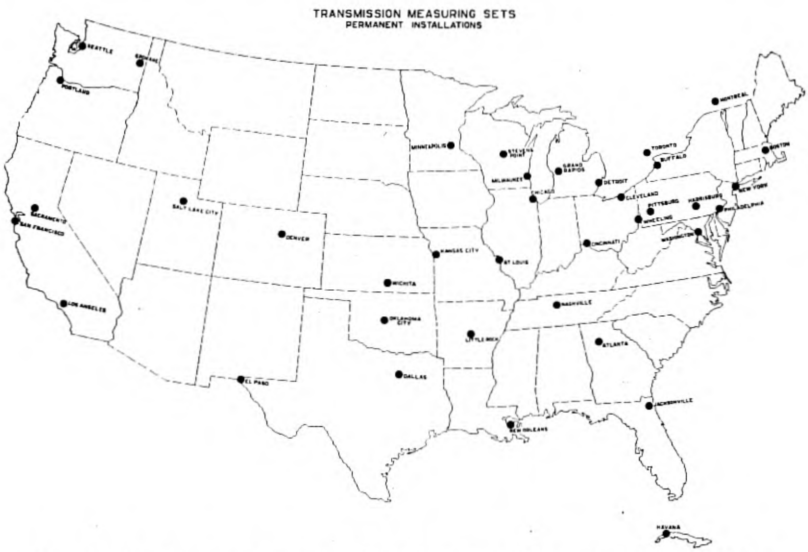


Fig. 3—Map Showing Locations in Bell System of Permanent Transmission Measuring Sets

measuring current is modulated and demodulated in the same manner as voice currents under regular operating conditions and the measured equivalent, therefore, indicates the overall transmission efficiency.

The map of Fig. 3 shows the locations in the Bell System of transmission measuring sets of the general type described above. At a number of the larger toll centers, such as New York and Chicago, where the number of toll circuits to be tested require it, several transmission measuring sets are installed. There are now in operation between 40 and 50 of these sets, making it possible to test all of the longer and more important toll routes in the system. The shorter toll circuits radiating out from the large toll centers are also tested with these same sets. At the smaller offices where fixed transmission measuring sets are not warranted, the toll circuits which cannot be picked up by the larger offices are tested by portable transmission measuring sets of the 1-A or 3-A types in connection with other maintenance work.

One very essential requirement in carrying on a systematic testing program is to have records of the detailed makeup of the toll circuits which give both the circuit layouts and the equipment associated with the circuits. Such a record is valuable, not only in giving the maintenance forces a picture of the circuits and equipment which they are

TOLL CIRCUIT LAYOUT RECORD

CIRCUIT NO. .... A .... B .... EQUIVALENT COMPUTED ...... MEASURED ...... | CIRCUIT ORDER .... ITEM .... DATE IN SERVICE .... CARD ISSUE NO. .... DATE ....

CONTROL OFFICE .... CLASSIFICATION ....

FROM 1	TO 2	CABLE OR LINE 3	PAIRS OR PINS 4	SIZE OF WIRE 5	LOADING 6	LENGTH 7	EQUIV. 8	REPEATING COILS ON SIDE 9	ON PHYS. 10	ON PHAN. 11	FOR SX 12	CX 13	RINGER 14	OTHER 15	TOTAL LOSS (COLUMNS 9 TO 15) 16
C															
D															
E															
F															
G															
H															
J															
K															
L															
M															
N															
P															
Q															
R															
S							TOTAL								TOTAL	

STATION 17	TELEPHONE REPEATER DATA			TOWARD A						TOWARD B						RINGING ON 2-W SIDE OF 4-W TERM. EQ. 33	DISTRIBUTION					
	CLASS 18	RINGING 19	IMP. 20	GAIN OR STEP EST. 21	GAIN OR STEP MEAS. 22	NET 23	S.O.S. 24	SINGING POINT EST. 25	SINGING POINT MEAS. 26	GAIN OR STEP EST. 27	GAIN OR STEP MEAS. 28	NET 29	S.O.S. 30	SINGING POINT EST. 31	SINGING POINT MEAS. 32		OFFICES	DIV.	DIST	TSO.	TMS	RPTR
T																						
U																						
V																						
W																						
Z																						

Fig. 4—Sample of a Toll Circuit Layout Record Card

testing, but it also furnishes a means for establishing the transmission standards to which they should work. When transmission tests indicate trouble, this record becomes of particular service in locating and clearing the cause.

Fig. 4 shows a sample of the type of toll circuit layout record card which has proven very satisfactory and is now generally used in the Bell System.

Telephone Repeater and Carrier Maintenance. Voice frequency telephone repeaters were discussed in a paper by Messrs. Gherardi and Jewett[3] and carrier systems in a paper by Messrs. Colpitts and Blackwell.[4] The various arrangements of amplifiers to provide for telephone repeater and for carrier operation as described in these papers make up integral parts of toll circuits and introduce elements

[3] Gherardi and Jewett, "Telephone Repeaters," *Transactions of A. I. E. E.*, 1919, Vol. XXVIII, part 2, pps. 1287 to 1345.

[4] Colpitts and Blackwell, "Carrier Current Telephony and Telegraphy," *Transactions of A. I. E. E.*, 1921, Vol. XL, pps. 205 to 300.

in the circuits which have to be given particular local attention in maintaining the overall transmission efficiency. Since both telephone repeaters and carrier employ the same types of vacuum tubes with very similar arrangements for power supply, the maintenance requirements for the two are much the same. The chief items to be observed in both carrier and repeater maintenance are that the gains specified to give a desired overall transmission equivalent be

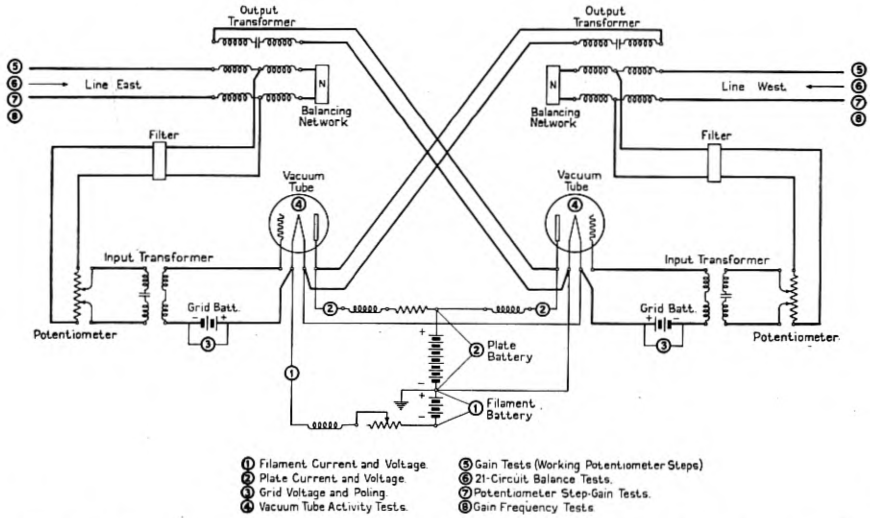
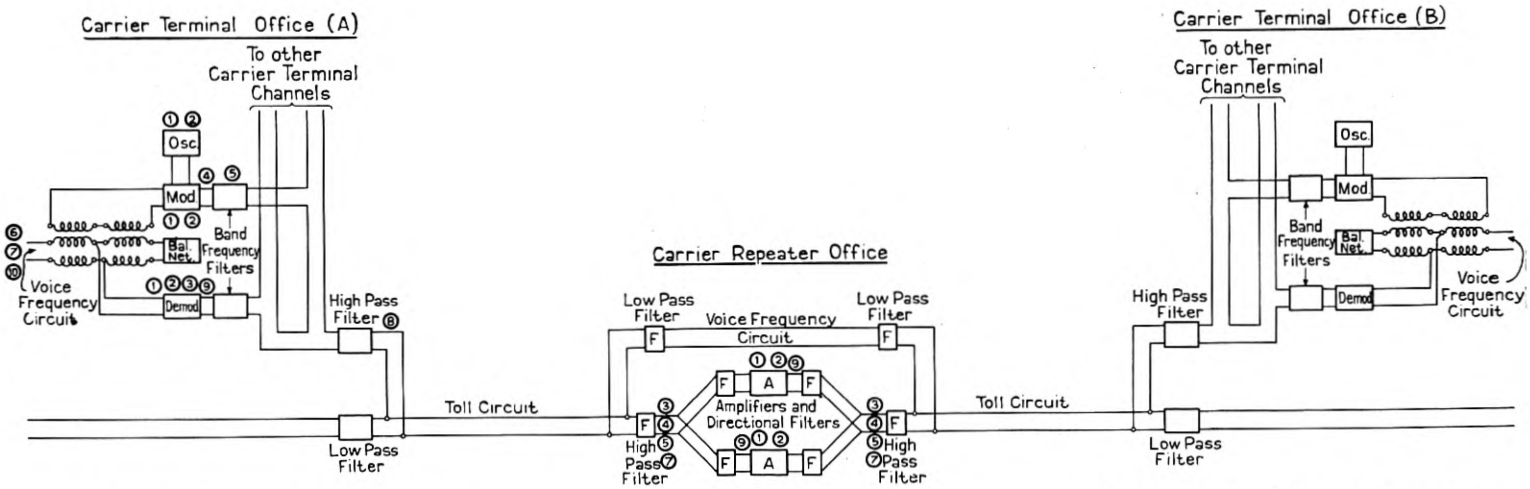


Fig. 5—Schematic Diagram of a 22-Type Telephone Repeater Showing Important Local Transmission Maintenance Tests

kept as constant as possible, that these gains remain fairly uniform within the range of frequencies involved, and that conditions do not exist which will disturb the overall balance between the circuits and networks sufficiently to cause poor quality of transmission.

Considering telephone repeater maintenance, Fig. 5 shows a schematic diagram of a 22 type repeater and indicates the important tests which are made locally to insure that the apparatus is functioning in a satisfactory manner as a part of a toll circuit. The numbers applied to the different tests listed in the figure show approximately the points in the repeater circuit at which the tests are made, the purposes of the tests being evident from their names.

When carrier operation is applied to toll circuits, an additional transmission system is introduced involving the use of currents of higher frequencies than those in the voice range. From a main-



Carrier Terminal Tests

1 Filament, Plate and Grid Battery Tests.
2 Vacuum Tube Activity Tests.
3 Channel Rectified Received Current Tests.
4 Modulator Output.
5 Modulator Band Filter Output.
6 Channel Loop Gain Tests.
7 21 Circuit Balance Tests on Voice Frequency Circuits.

Overall Tests of Complete Carrier System

8 Tests of total Carrier Output Current into Toll Circuit.
9 Tests of Carrier Current at Repeater Outputs and finally Rectified Received Current at Distant Terminal.
10 Overall Transmission Tests.

Carrier Repeater Tests

1 Filament, Plate and Grid Battery Tests
2 Vacuum Tube Activity Tests
3 Gain Tests.
4 Potentiometer Step-Gain Tests.
5 Gain-Frequency Tests.
6 Check of Frequency of Test Oscillator (not shown in figure.)
7 High Frequency Singing Tests.

9 Output Current on Overall Test of System.

Carrier Terminal Tests similar to those listed for Office (A)

Fig. 6—Schematic Diagram of a Carrier Telephone System Showing Important Transmission Maintenance Tests for Carrier Repeaters, Carrier Terminals and Overall

tenance standpoint this means that certain additional testing methods must be employed which will insure the proper generation and transmission of the carrier currents and that the modulation and demodulation of the voice frequency currents is accomplished without distortion or excess loss in overall transmission.

To give a general picture of the more important features involved in the transmission maintenance of carrier systems, Fig. 6 shows a schematic diagram of a carrier layout having one carrier repeater. The particular arrangement shown is for the type *B* system described by Messrs. Colpitts and Blackwell,[4] although the same general maintenance considerations apply to any of the present systems. It will be noted that three series of tests are required, one for the carrier repeaters, one for the carrier terminals and one for the system as a whole. The nature of these various tests and the approximate points in the carrier system where they are applied will be evident from the names and numbers used in the figure.

For both telephone repeaters and carrier systems, provision is made in the regular testing equipment so that the tests can be very quickly applied both as a routine proposition and also when required for trouble location.

Transmission Tests on Exchange Area Circuits

The transmission conditions in the exchange area plant are important not only from the standpoint of insuring good local service but also to insure good toll service, since the local plant forms the terminals of toll connections. The exchange or local plant offers a somewhat different transmission maintenance problem than the toll plant, particularly with respect to the routine testing procedures which must be followed to insure satisfactory transmission. This will be evident when it is considered that in each city and town a complete telephone system is in operation which involves the use of a large number of circuits of various types. There are also in use three general types of telephone switching equipments; manual, panel machine switching, and step-by-step machine switching, and in certain cities combinations of these equipments. It is estimated that at the present time in the Bell System there are in the neighborhood of two and one-half million exchange area circuits, exclusive of subscribers' lines, involving equipment other than contacts and wiring which may directly affect the transmission of speech.

The general classes of exchange area circuits in both manual and machine switching offices, important from a transmission maintenance standpoint, are listed in Table I. The operating features of manual telephone systems are generally well known as are also the features of step-by-step machine switching systems, both having been in use for many years. The panel machine switching system which is a relatively recent development was described in a paper by Messrs. Craft, Morehouse and Charlesworth.[5]

TABLE I

Classification of Circuits in the Exchange Area Plant Important from a Transmission Maintenance Standpoint

MANUAL OFFICES

Local Switchboards	P. B. X. Switchboards	Toll Switchboards	Toll Testboards
Cord circuits	Cord circuits	Cord circuits	Composite set circuits
Operators' circuits	Operators' circuits	Operators' circuits	Composite ringer circuits
Trunk circuits	Trunk circuits	Trunk circuits	Phantom & simplex circuits
Miscl. circuits	Miscl. circuits	Miscl. circuits	Miscl. circuits

Subscribers' loops and sets
Operators' telephone sets

MACHINE SWITCHING OFFICES

Panel	Step by Step
District selectors	Connectors
Incoming selectors	Toll selectors
Trunk circuits	Trunk circuits
Miscl. circuits	Miscl. circuits

Subscribers' loops and sets
Operators' telephone sets for
Special service positions

General classes of exchange area circuits involving equipment other than contacts and wiring which affect telephone transmission.

While it may appear at first hand from the above discussion that transmission testing in the exchange plant is a complicated and expensive matter, this has not proven to be the case. It has been found by experience that the systematic use of transmission measuring sets, following the testing methods which have been developed provides a means for periodically checking transmission conditions with a relatively small amount of testing apparatus and with a small maintenance force. All of the transmission circuits exclusive of subscribers' lines in a 10,000-line central office, either manual or machine switching, can, for example, be completely tested by two men in a

[5] Craft, Morehouse and Charlesworth, "Machine Switching Telephone System for Large Metropolitan Areas," *Journal of the A. I. E. E.*, April, 1923.

period of from two to four weeks, (five and one-half 8-hour days per week assumed) any trouble found being cleared as the testing work is done. The maintenance of the subscribers' lines is not included in this work since it is taken care of by other methods as outlined later.

In order to give a general picture of the application of transmission testing in the exchange telephone plant, a brief discussion of the methods employed in both manual and machine switching systems is given below. In either system the loop method of testing proves

Fig. 7—Illustration of a 3-A Transmission Measuring Set Being Operated in a Manual Office

most satisfactory, that is, one measuring set is used and where both terminals of a circuit are available as in cord circuits, a loop test through the circuit is made. In testing trunk circuits two trunks are looped together at their distant terminals and a measurement made on the two combined.

Transmission Tests on Manual Exchange Area Circuits. In central office, P. B. X. and toll switchboards, the cord circuits and associated operators' circuits are tested by using a portable transmission measuring set, moving this along the boards as required to pick up the cords. Fig. 7 shows a 3-A transmission measuring set being operated at an

A switchboard position. The cords are picked up and plugged directly into the set as shown and measurements made of the loss of both the cord and operator's circuits. Trunk circuit tests are made at the switchboards in the same manner as previously described for loop transmission tests on toll circuits, portable measuring sets such as shown in Fig. 7 generally being employed for this work. Operators' sets are inspected periodically and transmitter and receiver efficiency testing methods are under field trial which provide a means for testing these instruments in central offices. The miscellaneous transmission circuits in an office are tested at the points where they can be most conveniently picked up. The tests on toll test board circuits are made at this board and involve chiefly loop tests on the equipment associated with the toll circuits in the office and tests on the toll line circuits between the toll testboard and toll switchboard.

Transmission Tests on Machine Switching Circuits. The transmission circuits in panel machine switching systems are identical to those in manual systems, while these circuits in step-by-step systems are of a different design but essentially the same as far as transmission losses are concerned. Transmission tests on machine switching circuits are similar to those on manual circuits but involve special methods for picking up the circuits and holding them while the measurements are made. The standard types of transmission measuring sets are used in this work in conjunction with the regular testing equipment provided in the machine switching offices and the methods which have been developed offer a quick and convenient means for making the tests. In manual offices the circuits terminate in jacks or plugs at switchboards where they are readily accessible. In machine switching systems, provision is made for terminating the circuits in jacks at test desks or frames where they can be picked up by patching cords and tested as conveniently as the corresponding types of circuits in manual offices. Machine switching systems offer an important advantage in transmission testing work, particularly in trunk testing, in that the circuits to be tested can be looped automatically by the use of dials or selector test sets, thereby doing away with the necessity for having someone at the distant office complete the loops manually.

In panel machine switching offices the circuits involving transmission equipment corresponding to cord circuits are the "district" and "incoming" selectors. These are tested by setting up the transmission measuring set at the district or incoming frames and connecting the set to the test jacks associated with the circuits. Tests on trunks between manual and panel machine switching offices where both systems are in operation in the same exchange area are generally made from the manual office, the loops being dialed from the *A* switchboard, while trunks between two machine switching offices are tested from the outgoing end of the trunks.

Fig. 8 shows a 3-A transmission measuring set as used in a machine switching office ready for making tests on district selectors. To

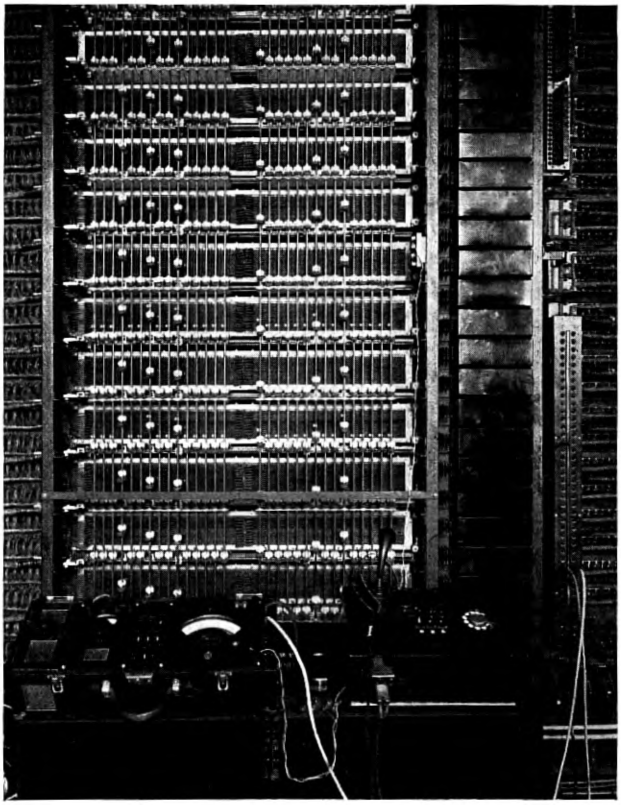


Fig. 8—Illustration of a 3-A Transmission Measuring Set, Set up in a Panel Machine Switching Office for Testing District Selectors

illustrate the general method of testing panel machine switching circuits, the upper diagram of Fig. 9 shows the schematic arrangement for measuring trunks between two panel machine switching offices. The transmission measuring set is located at office A, and connection made to the outgoing end of the trunks to office B through the test jacks at the trouble desk. A standard selector test set used in local maintenance work and a high impedance holding coil are also connected to the trunks through the measuring set, these being used to establish the loop and hold this loop while the tests are made. At office B two spare multiple circuits are cross-connected at the main distributing frame. Any two trunks in the group can then be auto-

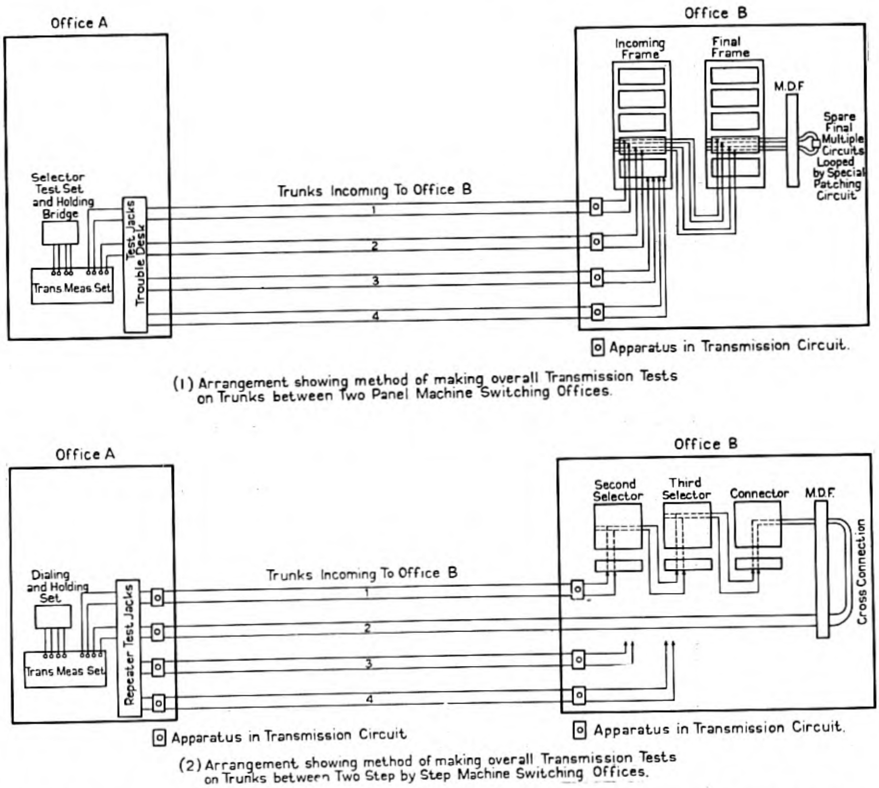


Fig. 9—Schematic Diagrams Showing Methods of Making Transmission Tests on (A) Trunks Between Panel Machine Switching Offices and (B) Between Step by Step Machine Switching Offices

matically looped together at office B by the use of the selector test set which functions to connect the trunks to the two spare multiple circuits previously cross-connected at office B.

In step-by-step machine switching offices the circuits involving transmission equipment corresponding to cord circuits are the connectors. Each connector is provided with a test jack through which connection can be made to a transmission measuring set and the loop completed over a test trunk by dialing. Local selectors do not contain any equipment other than contacts and wiring in the transmission circuits but these can be tested in the same manner as connectors if it is desired to check the wiping contacts and wiring. Toll selectors which involve equipment in the transmission circuit can also be tested in the same manner as connectors. Trunks between manual and machine switching offices can be most conveniently tested from the manual office, the trunk loops being established directly by dialing.

To illustrate the general method of testing step-by-step machine switching circuits, the lower diagram of Fig. 9 shows the schematic circuit arrangement for testing trunks between two machine switching offices. The transmission measuring set is located at office A in a position so that it can be patched to the outgoing trunk repeater test jacks and an arrangement for dialing and holding is connected to the trunks through the measuring set. At office B the apparatus in one trunk is disconnected and this trunk used as a test trunk by cross-connecting it at the main distributing frame to a spare subscriber's multiple terminal. All trunks in the group can then be tested by dialing over them, from office A, the number of this spare terminal at office B which automatically loops them back over the test trunk.

Maintenance of Subscribers' Lines and Stations. The circuits making up subscribers' lines from switchboard to instruments consist simply of pairs of conductors, almost always in cable, with the necessary protective devices. These can be checked by certain d-c. tests described in a recent paper.[6] Equipment is also provided in local test boards for use in making talking transmission tests between the station and the test boards. Accurate machine methods for determining the efficiency of transmitters and receivers have been developed for testing new instruments and instruments returned from service.

General Scheme of Testing Exchange Area Circuits. The plan being followed in the Bell System for systematically checking the transmission conditions of exchange area circuits is to have all offices tested periodically by men equipped with portable transmission measuring sets who travel from office to office. It has been found by experience that after an office has once been tested and any transmission troubles eliminated, it is only necessary thereafter to make transmission tests at infrequent intervals, these subsequent tests serving primarily as a check on the local maintenance conditions.

[6] W. H. Harden, "Electrical Tests and Their Applications in the Maintenance of Telephone Transmission," *Bell System Technical Journal*, July, 1924.

With a testing plan of this kind large areas can be covered by a small traveling force with a small amount of testing equipment. This results in a very economical transmission testing program while at the same time insuring that transmission conditions are maintained satisfactorily.

Fig. 10 shows a typical transmission testing team layout. The team is equipped with an automobile which proves an economical means of transportation between offices and exchange areas and

Fig. 10—Illustration of a Typical Transmission Testing Team Layout

provides a convenient method for carrying the testing equipment. During transportation this equipment is packed in padded trunks which insures against injury. In this particular case the equipment includes, in addition to transmission testing sets and oscillators, other apparatus such as a wheatstone bridge, crosstalk set and noise measuring set so that other maintenance work may be done in connection with transmission testing whenever this is desired.

Results Accomplished

The results accomplished in transmission maintenance work can best be appreciated by considering the kinds of troubles which adversely affect transmission and which can be detected and eliminated by routine testing methods. Consideration is first given to the general causes of troubles which are detrimental to both toll and local transmission, and later the features in this connection more particularly identified with telephone repeaters and carrier systems are discussed.

The different classes of circuits given in Table I are made up of various combinations of the following individual parts:

Repeating Coils	Plugs	
Retardation Coils	Jacks	
Relays	Keys	
Condensers	Heat Coils	
Resistances	Carbons	
Auto-Transformers	Wiring	Switchboard to M. D. F.
Induction Coils		Cross-connection
Loading Coils		Outside
Cords	Transmitters	
	Receivers	

The above parts are combined in various ways to make up the complete operating circuits such as cord circuits, operators' circuits, trunk circuits, etc. Each complete circuit causes a definite normal loss to telephone transmission which must be taken account of in designing the plant to meet the various service requirements. If, however, any of the parts used are defective, if the wrong combinations of parts are used, or if the installation work is not correctly done, excess transmission losses will result which may very seriously affect the transmission when the particular circuits involved are employed in an overall connection.

Classification of Common Types of Troubles. An analysis of a large amount of transmission testing data has made it possible to develop a definite trouble classification which is particularly helpful in transmission maintenance work and which permits the most efficient use of the results in eliminating transmission troubles. Experience has shown that the troubles found can be divided into two general classes, A—troubles which can be detected either by simple d-c. or a-c. tests in connection with the regular day-by-day maintenance work or by transmission measuring sets, and B—troubles which can be detected most readily by transmission measuring sets. The most important troubles in the above classes are as follows:

Class A	Class B
Opens	Electrical Defects
Grounds	Incorrect Wiring
Crosses	Wrong Type of Equipment
Cutouts	Missing Equipment
	High Resistance
	Low Insulation

If, in making transmission tests in a central office, a high percentage of Class A troubles is found the remedy is generally to instigate more rigid local maintenance routines paying particular attention to the type of circuits in which the troubles are located. The percentage of Class B troubles is not as a rule as high as the Class A troubles and experience has shown that when Class B troubles are once eliminated by transmission testing methods only infrequent subsequent tests are required to take care of any additional troubles of this class which may get into the plant.

In determining what constitutes an excess loss, the value of the transmission as well as the practical design and manufacturing considerations to meet operating limits are taken account of. An excess gain is also considered as a trouble on circuits equipped with amplifiers, since this may produce poor quality of transmission which is likely to be more detrimental to service than an excess loss. The value of transmission based on economical design considerations varies, depending on the first cost and annual charge of the particular types of circuits involved. A gain of one TU in the toll plant is generally worth more, for example, than one in the local plant, since it costs more to provide. In transmission maintenance work the cost of making transmission tests and clearing trouble is balanced against the value of the transmission gained for the purpose of establishing economical transmission limits to work to.

Specific Examples of Common Troubles Found and Their Effect on Transmission. Certain kinds of troubles which are detected by transmission measuring sets do not cause excess losses which can be quantitatively measured. Such troubles are, however, readily detected by "ear balance" transmission measuring sets in that they cause noise or scratches and by the "meter balance" sets from fluctuations of the needle of the indicating meter. The most common trouble of this kind is due to cutouts or opens which may be caused by dirty connections, loose connections, improper key and relay adjustments, etc. While not causing a quantitative value of excess loss, this class of trouble is very detrimental to transmission and more serious in many instances than fixed excess losses. Indeterminate troubles of this nature are given an arbitrary excess loss value based on experience.

Considering troubles which give definite losses, the most common kinds are caused by electrical defects in equipment, incorrect wiring of equipment in circuits and wrong types of equipment. The other classes of troubles, such as crosses, high resistances, and low insulation, also generally give measurable excess losses but these are not as common in the plant, since troubles of this nature are more likely to affect the signaling and operation of the circuits and are, therefore, eliminated by the regular maintenance work. Missing equipment will in certain cases cause a gain in transmission but affects the circuits adversely in other ways.

Typical examples of common troubles, with the excess losses which they cause, are given in the following table:

Type of Circuit and Equipment	Cause of Trouble	Approximate Excess Transmission Loss[7]
Repeating coils in cords, incoming trunk circuits, selectors, toll connectors	Electrical defects (Generally short circuited turns)	1.5 to 5.0 TU
	Incorrect wiring (Generally reversed windings)	2.0 to 13.0 TU
Supervisory relays in "A" cord circuits	Electrical defects (Open non-inductive winding)	About 2.5 TU
Bridged retardation coils or relays in toll cord circuits, composite sets, connectors and step-by-step repeaters	Electrical defects (Generally short circuited turns)	1.0 to 5.0 TU
Repeating coils on loaded toll switching trunks	Wrong type of equipment, incorrect wiring	1.0 to 4.0 TU
Induction coils in operators' telephone sets	Electrical defects. Incorrect wiring	1.0 to 13.0 TU

There are, of course, many other specific types of troubles detected by transmission tests which give definite quantitative losses but the above will serve to illustrate the value of this testing work in eliminating excess losses in a telephone plant.

Maintenance Features Peculiar to Telephone Repeaters and Carrier Systems. The same classification of troubles discussed above applies to repeaters and carrier systems. Amplifier equipment, however, employs certain features which are not common to the more simple telephone circuits and some of the troubles which may occur if the proper maintenance procedures are not followed will seriously affect service. It is for this reason that repeater and carrier installations are provided with special testing equipment which is always available for use either in routine maintenance or in locating and clearing any troubles which may occur in service. Automatic regulating devices are also provided wherever this is practicable in order to reduce to a minimum the amount of manual regulation and maintenance.

[7] W. H. Martin, "The Transmission Unit," *Journal of the A. I. E. E.*, June, 1924; *B. S. T. J.*, Vol. III, p. 400, 1924; C. W. Smith, "Practical Application of Transmission Unit," *B. S. T. J.*, Vol. III, p. 409, 1924.

The important elements in both repeaters and carriers which may directly affect transmission or cause service troubles in other ways are as follows:

Filament Batteries	Potentiometers
Plate Batteries	Filters
Grid Batteries	Transmission Equalizers
Vacuum Tubes	Signaling Equipment
Balancing Equipment	Patching Arrangements

The tests outlined in the main body of the paper aim to insure that the above essential parts of repeater and carrier circuits are functioning properly and that the equipment as a whole is giving the desired results in overall transmission efficiency.

Conclusion

The above discussion of testing methods and the results accomplished indicate how a comprehensive and economical transmission maintenance program can be applied to a telephone plant to check the volume efficiency of the circuits against the established standards. Consideration is continually being given to new testing methods and their applications in order that further improvements in service may be effected and increased economies in testing taken advantage of.

APPENDIX A

Principles of Testing Methods Closely Associated with Transmission Efficiency Tests

Tests of volume efficiency often need to be supplemented by other methods of testing in transmission maintenance work. Transmission efficiency both as regards volume and quality may be seriously affected by noise or crosstalk, and tests for any conditions of this kind are therefore important in maintenance work. Furthermore when efficiency tests show excess losses or unsatisfactory circuit conditions other testing methods prove very valuable in locating the cause.

To illustrate this phase of transmission maintenance the principles of some of the more important testing methods are briefly described below. Two of the tests employ a method very similar to loop transmission testing while others employ the well known "null" method. A special method employing three winding transformers and amplifiers widely used to determine impedance balance conditions between lines and networks is also described. Several methods which involve simply current and voltage measurements have been mentioned in this paper but these are generally well known and therefore require no detailed description.

1. Measurements of Crosstalk

In the circuit shown in Fig. 11, if a-c. power is supplied to a circuit known as the "disturbing" circuit and unbalances exist between this circuit and a second known as the "disturbed" circuit, power will be transferred from one circuit to the other causing crosstalk in the second. A definite power transmission loss therefore takes place between the two circuits which can be measured by a loop transmission test similar to the efficiency tests described in the main body of the paper. An adjustable shunt called a "crosstalk meter" cali-

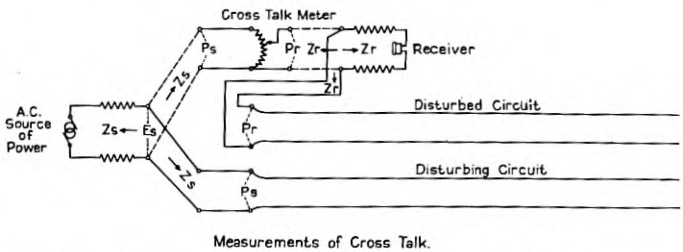


Fig. 11—Diagram Showing Principles of Crosstalk Measurements

brated in either TU or in crosstalk units is substituted for the two circuits. With the same power supplied alternately to both the "disturbing" circuit and the meter and with the sending and receiving end impedance conditions as shown, the meter shunt is adjusted until, in the opinion of the observer, the annoyance produced by the tone in the receiver is judged to be equal for the two conditions. The reading of the shunt if there was no distortion of the line crosstalk currents would then give the volume of crosstalk which could be expressed in TU as $10 \log_{10} P_r/P_s$ similar to loop transmission testing. However, this relation only holds approximately in practise since the line crosstalk measured is produced by various currents having different phase relations and a certain amount of distortion therefore occurs. The commercial form of crosstalk set now used is equipped to give the approximate impedance relations required and also provides a feature for eliminating the effect of line noise except in the case of one type of measurement which is made on long cable circuits. For practical reasons the results are generally expressed in crosstalk units rather than TU.

2. Measurements of Noise

The common method of measuring noise in a telephone circuit is shown in the diagram of Fig. 12. In this test an artificial noise current produced by a generator of constant power P_s called a "noise standard" is substituted for the line noise current. If the two noise currents were exactly alike as regards wave shape and the relative magnitude of the frequencies involved they would produce the same tone in the receiver and their volumes could be made equal by adjustment of the noise shunt. The power ratio, P_r/P_s, as indicated

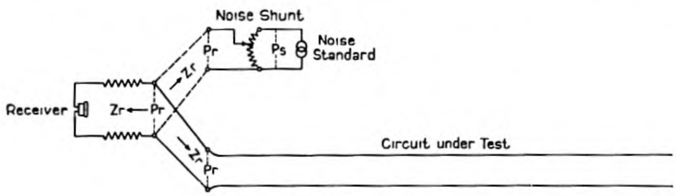
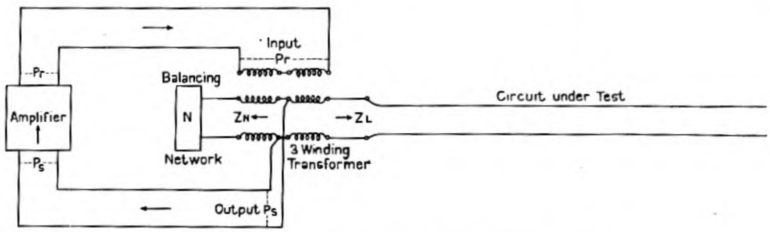


Fig. 12—Diagram Showing Principles of Noise Measurements

by the shunt, would then give a measure of the line noise in terms of the noise standard. This condition, however, is not met with in practise due to differences in wave shape of the two noise currents. For this reason noise measurements are made by adjusting the noise shunt until the interfering effects of the noise on the line and from the shunt are judged to be the same for which condition the power supplied to the receiving network by the noise standard is not necessarily the same as that supplied by the line. The receiving end impedances however, are kept as nearly alike as practicable to prevent reflection losses.

3. Measurements of Line-Network Balance (21-Circuit Balance Test)

The testing arrangement of Fig. 13 shows the principle of the 21-circuit balance test referred to in the main body of the paper in connection with telephone repeater and carrier maintenance. In this test the gain of an amplifier calibrated in TU is used to compensate for the loss through a three winding transformer or output coil of a telephone repeater. If the impedances of the balancing network and line were exactly alike at all frequencies, *i.e.*, $Z_n = Z_L$, and no other unbalances existed in the circuit none of the power supplied by the amplifier to the input of the three-winding transformer would be transferred to the output, *i.e.*, the power ratio P_s/P_r would be infinity. However, this ideal condition cannot be produced in practise so that there is always a finite power loss between the input and output of the transformer which can be measured approximately by the gain of an amplifier calibrated in *TU*. An internal



Measurements of Impedance Balance Between Lines and Networks
(21-Circuit Tests on Telephone Repeaters and Carrier)

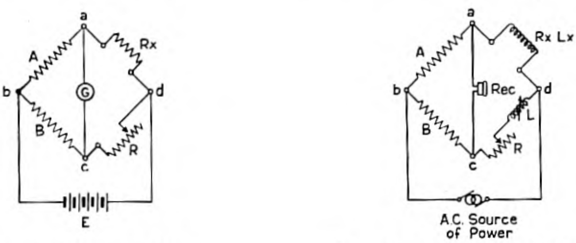
Fig. 13—Diagram Showing Principles of 21 Circuit Balance Tests

path for currents which may produce "singing" or a sustained tone is established if the gain of the amplifier P_r/P_s is greater than the loss P_s/P_r through the three-winding transformer. As unbalances between network and line become greater the loss through the three-winding transformer becomes less thereby requiring less gain in the amplifier to produce a "singing" condition. It should be noted in this connection that to produce the condition described above exactly, the current received around the "singing" path must be in phase with the starting current. In practise this condition obtains sufficiently accurately so that the gain of the amplifier required to produce "singing" gives an approximate measure of the impedance balance between line and network.

4. Measurements of Resistance, Reactance and Impedance

Diagram (a) of Fig. 14 shows the wheatstone bridge circuit for d-c. resistance measurements. It is unnecessary to describe the well known principles of this bridge but mention is made of it here in view of its importance and use in telephone maintenance work. It supplies an indispensable method of measurement for certain trouble locations, such as crosses and grounds and embodies the fundamental principles of all null tests.

Diagram (b) of Fig. 14 gives a bridge circuit for measuring impedance, the particular arrangement shown being for measurements of impedances having inductive reactance. The bridge measurements express impedance in terms of its resistance component and equivalent inductance or capacity. In measuring an impedance having inductive reactance at any frequency, f, for example, a balance gives $R=R_x$ and $L=L_x$. At the frequency f, the effective resistance is given directly by the value of R and the reactance by the relation, $2\,\pi f L$.



(a) Direct Current Resistance. (b) Impedance having Inductive Reactance.

Null Method of Measuring Resistance, Reactance and Impedance.

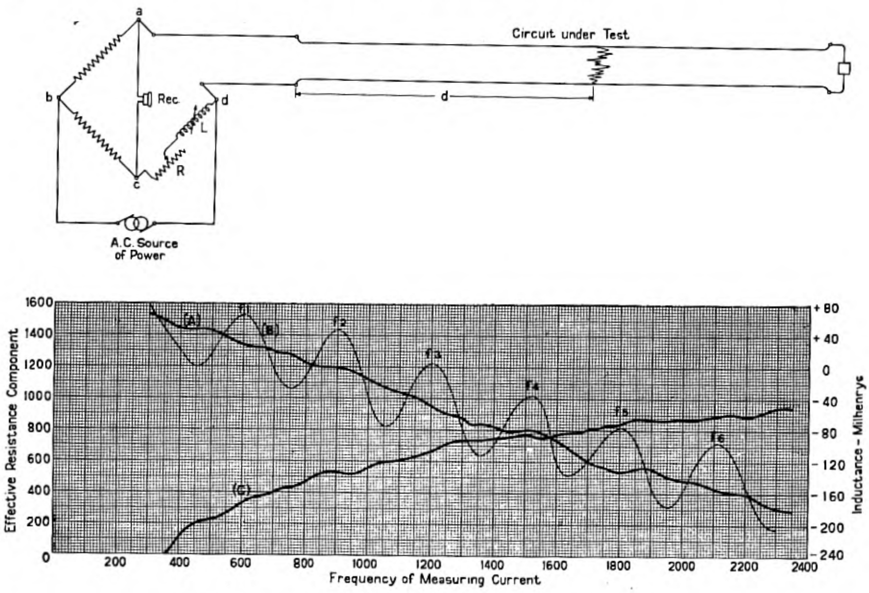
Fig. 14—Diagrams Showing Principles of Null Methods for Measuring Resistance, Reactance and Impedance

The impedance is the vectorial sum of these two or $\sqrt{R^2+(2\,\pi f L)^2}$. In maintenance work involving impedance measurements as will be noted in the next testing method described, the effective resistance component and the equivalent inductance are generally used directly without combining.

5. Measurements of Line Impedance and Location of Impedance Irregularities

Fig. 15 shows a telephone circuit connected to a bridge and terminated at its distant end in characteristic impedance. If the circuit has approximately uniform impedance throughout its length the resistance and equivalent inductance curves of this impedance within a range of frequencies will be fairly smooth as indicated by A and C of the figure. The curves are not perfectly smooth since it is not practicable to construct the line for perfect impedance uniformity. If at some point in the circuit an irregularity is present such as an omitted loading coil, an inserted length of line of different construction, etc., which changes the impedance, this will produce a periodic change in the resistance and inductance curves A and C such as shown by Curve B. Curve C will be changed in the same way as Curve A but for simplification this is not shown on the diagram.

The change in impedance in the circuit reflects some of the current sent out back to the sending end where it adds to or subtracts from the sending current depending on the phase relations of the two currents at any particular frequency. Since impedance equals E/I its value changes as the value of I changes. This is made use of in line impedance measuring work to give a location of impedance irregularities which may exist somewhere in the line.



Null Method of Measuring Line Impedance and Locating Impedance Irregularities

Fig. 15—Diagram and Impedance Curves Showing Principles of Line Impedance Measurements by Null Method and Location of Impedance Irregularities

Referring to Fig. 15, let d equal the distance in miles to an impedance irregularity and f, one frequency at which the resistance component of the impedance is a maximum. The next maximum point will occur at a frequency f_2 such that as the frequency has been increased, one complete wave length is added in the distance traveled by the reflected current. Maximum points at f_3, f_4, etc., occur in the same way as the frequency is increased. Considering the two values f_1 and f_2 let

V = velocity of current in miles per second

W_1 = wave length at frequency f_1

W_2 = wave length at frequency f_2

N = number of wave lengths in distance traveled by reflected current or $2\,d$.

At frequency f_1 then,

$$N=\frac{2d}{W_1}$$

and at f_2,

$$N+1=\frac{2d}{W_2}$$

also at f_1,

$$W_1=V/f_1$$

and at f_2,

$$W_2=V/f_2$$

Substituting above

$$N=\frac{2df_1}{V} \text{ and}$$

$$N+1=\frac{2df_2}{V}$$

Subtracting,

$$1=\frac{2df_2}{V}-\frac{2df_1}{V} \text{ or}$$

$$d=\frac{V}{2(f_2-f_1)}$$

which is the distance in miles from the sending end of the circuit to the point of impedance irregularity. The velocity of propogation V is not exactly constant within the entire frequency range but does not vary sufficiently to materially effect the accuracy of impedance trouble locations by this method.

Mutual Inductance in Wave Filters with an Introduction on Filter Design

By K. S. JOHNSON and T. E. SHEA

PART I

General Principles of Wave Filter Design

Principles of Generalized Dissymmetrical Networks. We shall consider first the impedance and propagation characteristics of certain generalized networks. It can be shown that *any passive network having one pair of input and one pair of output terminals may, at any frequency, be completely and adequately represented by an equivalent T or π network.*[1] The impedance and propagation characteristics of any such network may be expressed in terms of its equivalent T or π network. These characteristics are defined by (1) the *image impedances*, and (2) the *transfer constant*, the latter including the *attenuation constant*[2] and the *phase constant*.[2] In the case of a symmetrical network, the image impedances and the transfer constant are, respectively, the *iterative impedances* (or *characteristic impedances*) and the *propagation constant* employed by Campbell, Zobel, and others. The terms involved will be subsequently defined.

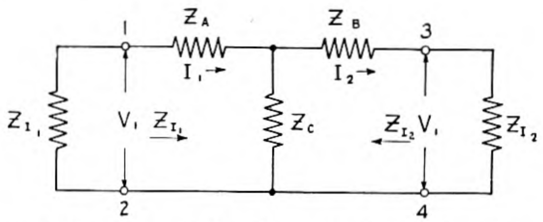


Fig. 1—Generalized Dissymmetrical T Network Connected to Impedances Equal to Its Image Impedances

Consider the dissymmetrical T network of Fig. 1. If the $3-4$ terminals of the T network are connected to an impedance Z_{I_2}, the

[1] Campbell, G. A., "Cisoidal Oscillations," *Transactions A. I. E. E.*, (1911), Vol. XXX, Part II, pp. 873–909.
The T and π networks referred to above are sometimes called *star* (Y) and *delta* (Δ) networks, respectively.

[2] The real and imaginary parts of the transfer constant have been called by Zobel, the *diminution constant* and the *angular constant*, respectively. (See Bibliography 13.)

impedance looking into the T network at the $1-2$ terminals will be

$$Z_{1-2}=Z_A+\frac{Z_C(Z_B+Z_{I_2})}{Z_C+Z_B+Z_{I_2}}. \tag{1}$$

Similarly, if the $1-2$ terminals of the T network are connected to an impedance Z_{I_1}, the impedance looking into the $3-4$ terminals of the T network will be

$$Z_{3-4}=Z_B+\frac{Z_C(Z_A+Z_{I_1})}{Z_C+Z_A+Z_{I_1}}. \tag{2}$$

If Z_{1-2} is equal to the terminal impedance Z_{I_1}, and if, similarly, Z_{3-4} is equal to the terminal impedance Z_{I_2}, the network will then be terminated in such a way that, at either junction ($1-2$ or $3-4$), the impedance in the two directions is the same. In other words, at each junction point, the impedance looking in one direction is the *image* of the impedance looking in the opposite direction. Under these conditions Z_{I_1} and Z_{I_2} are called the *image impedances* of the T network. If equations (1) and (2) are solved explicitly for Z_{I_1} and Z_{I_2}, the following expressions are obtained:

$$Z_{I_1}=\sqrt{\frac{(Z_A+Z_C)(Z_AZ_B+Z_AZ_C+Z_BZ_C)}{(Z_B+Z_C)}}, \tag{3}$$

$$Z_{I_2}=\sqrt{\frac{(Z_B+Z_C)(Z_AZ_B+Z_AZ_C+Z_BZ_C)}{(Z_A+Z_C)}}. \tag{4}$$

If Z_{oc} is the impedance looking into one end of the network with the distant end open-circuited, and if Z_{sc} is the corresponding impedance with the distant end short-circuited, it may be shown that the image impedance at either end of the network is the geometric mean of Z_{oc} and Z_{sc}. What is here termed the image impedance is, therefore, equivalent to what Kennelly has called the *surge impedance.*[3]

The propagation characteristics of a dissymmetrical network may be completely expressed in terms of the transfer constant. The transfer constant of any structure may be defined as one-half the natural logarithm of the vector ratio of the steady-state vector volt-amperes entering and leaving the network when the latter is terminated in its image impedances. The ratio is determined by dividing the value of the vector volt-amperes at the transmitting end of the network by the value of the vector volt-amperes at the receiving end.

[3] There is at present lack of common agreement as to the basis of definition of this term, and it is often defined upon the basis, not of open and short-circuit impedances, but of a uniform recurrent line (See A. I. E. E. Standardization Rule 12054, edition of 1922). The formulae derived by the two methods are not equivament in the case of dissymmetrical networks.

The real part of the transfer constant, that is, the *attenuation constant*, is expressed by the above definition in *napiers* or *hyperbolic radians* and the imaginary part, that is, the *phase constant*, is expressed in circular *radians*. The practical unit of attenuation here

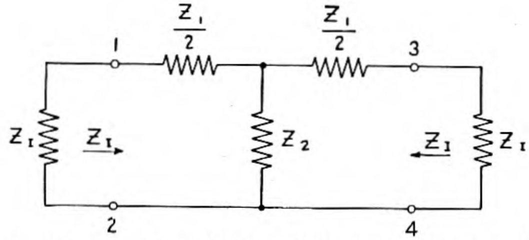


Fig. 2—Generalized Symmetrical T Network Connected to Impedances Equal to Its Image Impedances

used is the *transmission unit*[4] (1 $TU = .11513$ napier). It can be demonstrated that the transfer constant, Θ, of the T network shown in Fig. 1 is

$$\Theta = \tanh^{-1}\sqrt{\frac{Z_{sc}}{Z_{oc}}} = \tanh^{-1}\sqrt{\frac{Z_AZ_B+Z_AZ_C+Z_BZ_C}{(Z_A+Z_C)(Z_B+Z_C)}}$$
$$= \cosh^{-1}\sqrt{\frac{(Z_A+Z_C)(Z_B+Z_C)}{Z_C^2}}, \tag{5}$$

in which Z_{oc} and Z_{sc} are, as previously defined, the open and short-circuit impedances of the network. The ratio Z_{sc}/Z_{oc} is the same at both ends of any passive network.

Principles of Generalized Symmetrical Networks. Consider now the impedance and propagation characteristics of the generalized symmetrical structure shown in Fig. 2. On account of the symmetry of the structure, the image impedances at both ends are identical, and from equation (3) or (4) their value may be shown[5] to be

$$Z_I = \sqrt{Z_1Z_2\left(1+\frac{Z_1}{4Z_2}\right)}. \tag{6}$$

In the case of a symmetrical T structure, such as is shown in Fig. 2, the impedance Z_I is called the *mid-series image impedance.* The significance of this term will be evident, if the series-shunt type of

[4] W. H. Martin, "The Transmission Unit and Telephone Transmission Reference System," *Bell Sys. Tech. Jour.*, July, 1924; *Jour. A. I. E. E.*, Vol. 43, p. 504, 1924.

[5] Zobel, O. J., "Theory and Design of Uniform and Composite Electric Wave-Filters," *Bell Syst. Tech. Jour.*, Jan., 1923.

structure shown in Fig. 3 is regarded as made up of symmetrical T networks or sections, the junctions of which occur at the mid-points of the series arms.

Suppose now that the structure of Fig. 3 is considered to be made up of symmetrical π networks, or sections, each of which is represented

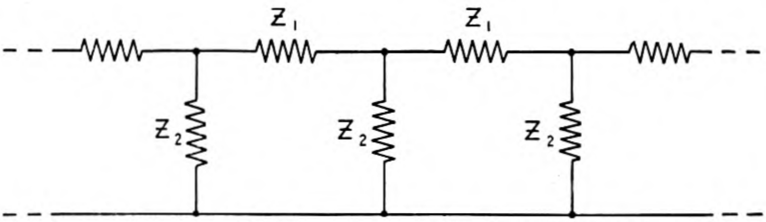


Fig. 3—Generalized Recurrent Series-Shunt Network

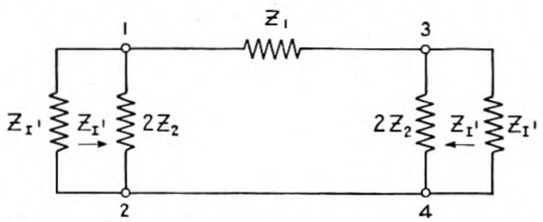


Fig. 4—Generalized Symmetrical π Network Connected to Impedances Equal to Its Image Impedances

as in Fig. 4. By methods similar to those employed for the T network of Fig. 2 it can be shown[5] that the image impedance of the generalized π network of Fig. 4 is given by

$$Z_{I'} = \sqrt{\frac{Z_1 Z_2}{1+\frac{Z_1}{4Z_2}}}. \tag{7}$$

In this symmetrical structure the image impedance is called the *mid-shunt image impedance.*

The image transfer constant of either a T or a π symmetrical structure is[5]

$$\Theta = A + jB = 2\ \sinh^{-1}\sqrt{\frac{Z_1}{4Z_2}} = \cosh^{-1}\left(1+\frac{Z_1}{2Z_2}\right). \tag{8}$$

In discussing the generalized networks of Figs. 1, 2 and 4, it has been assumed that the networks were terminated in their respective image impedances. In practical cases, filters must be designed to work between impedances which are, in general, not exactly equal to their image impedances at more than one or a few frequencies. For a generalized structure, such as that of Fig. 1, operating between a *sending-end impedance* Z_S and a *receiving-end impedance* Z_R, the current in Z_R, for an electromotive force acting in Z_S, is

$$I_R=\frac{E}{Z_S+Z_R}\times\frac{Z_S+Z_R}{\sqrt{4Z_SZ_R}}\times\frac{\sqrt{4Z_{I_1}Z_S}}{Z_{I_1}+Z_S}\times\frac{\sqrt{4Z_{I_2}Z_R}}{Z_{I_2}+Z_R}$$
$$\times\epsilon^{-\theta}\times\frac{1}{1-\frac{Z_{I_2}-Z_R}{Z_{I_2}+Z_R}\times\frac{Z_{I_1}-Z_S}{Z_{I_1}+Z_S}\times\epsilon^{-2\theta}}. \tag{9}$$

Since $E/(Z_S+Z_R)$ is the current ($I_{R'}$) which would flow if the generalized T network were not inserted in the circuit, the ratio of the received current, *with* and *without* the network in the circuit, may be expressed by the relation

$$\frac{I_R}{I_{R'}}=\left(\frac{Z_S+Z_R}{\sqrt{4Z_SZ_R}}\right)\left(\frac{\sqrt{4Z_{I_1}Z_S}}{Z_{I_1}+Z_S}\right)\left(\frac{\sqrt{4Z_{I_2}Z_R}}{Z_{I_2}+Z_R}\right)$$
$$\times\epsilon^{-\theta}\times\frac{1}{1-\left(\frac{Z_{I_2}-Z_R}{Z_{I_2}+Z_R}\right)\left(\frac{Z_{I_1}-Z_S}{Z_{I_1}+Z_S}\right)\epsilon^{-2\theta}}. \tag{10}$$

In general, the electromotive force does not act through a simple sending-end impedance Z_S but through some complex circuit. The current ratio ($I_R/I_{R'}$) will, however, be the same in either case. The principle underlying this fact is known as *Thévenin's Theorem.*[6]

The absolute magnitude of the current ratio, $|I_R/I_{R'}|$, is a measure of the *transmission loss* caused by the introduction of the network. The transmission loss may be expressed in terms of transmission units (TU) by aid of the following relation

$$TU=20\log_{10}\left|\frac{I_{R'}}{I_R}\right|. \tag{11}$$

Reference to equation (10) shows that the transmission loss caused by the introduction of any network is composed of five factors. The first three factors of this equation are all of the same general type with the exception that the first of the three is reciprocal in nature to the other two. These two latter factors have been called *reflection factors* and determine the *reflection losses* which exist between the impedances involved. The fourth factor is the *transfer factor* and expresses the current ratio which corresponds to the transfer con-

[6] Casper, W. L., "Telephone Transformers," *Transactions A. I. E. E.*, March, 1924, p. 4. Thévenin, M. L., "Sur un Nouveau Théorème d'Electricité Dynamique," *Comptes Rendus*, vol. 97, p. 159, 1883.

stant. The last factor has been called the *interaction factor*. The value of the reflection factor is evidently a function simply of the *ratio* of the impedances involved, while the absolute value of the transfer factor is ϵ^{-A} where A is the real portion of the transfer constant and hence is the attenuation constant. The value of the interaction factor is seen to be unity either when $Z_{I_2}=Z_R$ or when $Z_{I_1}=Z_S$. It also approaches unity if the value of Θ is sufficiently large.

In the case of a symmetrical structure, such as is shown in Fig. 2, or Fig. 4, $Z_{I_1}=Z_{I_2}=Z_I$ and equation (10) reduces to

$$\frac{I_R}{I_{R'}}=\left(\frac{Z_S+Z_R}{\sqrt{4Z_SZ_R}}\right)\left(\frac{\sqrt{4Z_IZ_S}}{Z_I+Z_S}\right)\left(\frac{\sqrt{4Z_IZ_R}}{Z_I+Z_R}\right)$$

$$\times\epsilon^{-\theta}\times\frac{1}{1-\left(\dfrac{Z_I-Z_R}{Z_I+Z_R}\right)\left(\dfrac{Z_I-Z_S}{Z_I+Z_S}\right)\epsilon^{-2\theta}}. \tag{12}$$

If the structure is symmetrical, and if, furthermore, the sending-end impedance Z_S is equal to the receiving-end impedance Z_R, equation (12) becomes

$$\frac{I_R}{I_{R'}}=\epsilon^{-\theta}\times\frac{4Z_IZ_R}{(Z_I+Z_R)^2}\times\frac{1}{1-\left(\dfrac{Z_I-Z_R}{Z_I+Z_R}\right)^2\epsilon^{-2\theta}}. \tag{13}$$

The preceding formulae make it possible to calculate rigorously the transmission loss caused by any network whose image impedances and transfer constant are both known. In the symmetrical case, if $Z_I=Z_S=Z_R$, the transmission loss is determined simply by the value of the attenuation constant. In general, in the attenuation range of frequencies, the value of Θ of a wave filter is relatively large and the interaction factor is substantially unity. Consequently, the transmission loss caused by any filter in its attenuation range is dependent practically only upon the value of the attenuation constant and the reflection losses between Z_S and Z_{I_1}, Z_R and Z_{I_2}, and Z_S and Z_R, respectively. Throughout most of the transmission range of a filter, its image impedances may be made very closely equal to the terminating impedances so that the transmission loss caused by the filter in this range is dependent simply upon its attenuation constant. In the intervening range, between the attenuated and the non-attenuated bands, the transfer factor, the reflection factors and the interaction factor must all be taken into account.[7]

[7] Zobel, O. J., "Transmission Characteristics of Electric Wave-Filters," *Bell Sys. Tech. Jour.*, Oct., 1924.

Impedance and Propagation Characteristics of Non-Dissipative Filters. If the series and shunt impedances of the structures shown in Figs. 2 and 4 are pure reactances, as they would be in the case of a non-dissipative filter, the ratio of the quantity $Z_1/4Z_2$ must be either a positive or negative numeric. It has been shown by Campbell[8] and others that the attenuation constant is zero, and that the structure freely transmits at all frequencies at which the ratio $Z_1/4Z_2$ lies between 0 and -1. Therefore, by plotting values of the ratio $Z_1/4Z_2$ it is possible to determine the attenuation characteristic of any symmetrical structure as a function of frequency.

In the transmission range, the phase constant of the symmetrical structure shown in Fig. 2 or Fig. 4, is

$$B=2\ \sin^{-1}\sqrt{\frac{-Z_1}{4Z_2}}. \tag{14}$$

Hence, the expression for the image transfer constant of either of the symmetrical structures shown in Fig. 2 or Fig. 4 is

$$\Theta=0+j\ 2\ \sin^{-1}\sqrt{\frac{-Z_1}{4Z_2}}. \tag{15}$$

In the attenuation region, $Z_1/4Z_2$ may be either negative or positive. If $Z_1/4Z_2$ is negative and is greater in absolute magnitude than unity, the attenuation constant is

$$A=2\ \cosh^{-1}\sqrt{\frac{-Z_1}{4Z_2}} \tag{16}$$

and the phase constant, or the imaginary component of the image transfer constant, is

$$B=(2K-1)\pi \tag{17}$$

where K is any integer. Hence,

$$\Theta=2\ \cosh^{-1}\sqrt{\frac{-Z_1}{4Z_2}}+j(2K-1)\pi. \tag{18}$$

From equation (8), when $Z_1/4Z_2$ is positive, the attenuation constant is

$$A=2\ \sinh^{-1}\sqrt{\frac{Z_1}{4Z_2}} \tag{19}$$

and the phase constant B is zero. Hence,

$$\Theta=2\ \sinh^{-1}\sqrt{\frac{Z_1}{4Z_2}}+j0. \tag{20}$$

[8] Campbell, G. A., "Physical Theory of the Electric Wave-Filter," *Bell Sys. Tech. Jour.*, Nov., 1922.

As a result of equations (18) and (20), in the attenuation range, the phase constant of a non-dissipative symmetrical filter section is always zero or an odd multiple of $\pm\pi$.

The *cut-off frequencies*, by which are meant the divisional frequencies which separate the transmission bands from the attenuation bands, must always occur when $Z_1/4Z_2=0$ or when $Z_1/4Z_2=-1$, since, for the transmission bands, $Z_1/4Z_2$ must lie between 0 and -1.

The general formulae for the image impedances of the symmetrical networks shown in Figs. 2 and 4 are equations (6) and (7), respectively. From these equations, the image impedances are pure resistances in the transmission range of a non-dissipative structure. In the attenuation range, however, the image impedances are pure reactances; the mid-series image impedance is a reactance having the same sign as Z_1, while the mid-shunt image impedance is a reactance having the same sign as Z_2. In these attenuation bands, the image impedances (pure reactances) have positive or negative signs depending upon whether they are increasing or decreasing with frequency. The order of magnitude of the image impedances may be found from Table I.

TABLE I

If the Value of $\lvert Z_1 \rvert$ is	And if the Value of $\lvert 4Z_2 \rvert$ is	Then the Mid-Series Image Impedance is	And the Mid-Shunt Image Impedance is
Zero	Zero	Zero	Zero
Zero	Finite	Zero †	Zero †
Zero	Infinite	Finite †	Finite †
Finite	Zero	Finite **	Zero **
Finite	Finite	Zero* or Finite	Infinite * or Finite
Finite	Infinite	Infinite †	Infinite †
Infinite	Zero	Infinite **	Zero **
Infinite	Finite	Infinite **	Finite **
Infinite	Infinite	Infinite	Infinite

* When both Z_1 and Z_2 are finite and $Z_1=-4Z_2$, the mid-series image impedance is zero and the mid-shunt image impedance is infinite.

† This condition gives a cut-off frequency.

** This condition results in infinite attenuation.

Types of Non-Dissipative Series-Shunt Sections Having Not More Than One Transmission Band or More Than One Attenuation Band. Since the series and shunt arms of a non-dissipative filter section may each be composed of any combination of pure reactances, it is possible to have an infinite number of types of filter sections. However, it is seldom desirable to employ filters having more than one transmission band or more than one attenuation band. Under these conditions, it is generally impracticable to employ more than four reactance elements in either of the arms of a section. Likewise, a total of six reactance elements in both the series and shunt arms is the maximum that can be economically employed.

Types of two-terminal reactance meshes having not more than four elements, are listed in Fig. 5. In Fig. 6, the corresponding frequency-

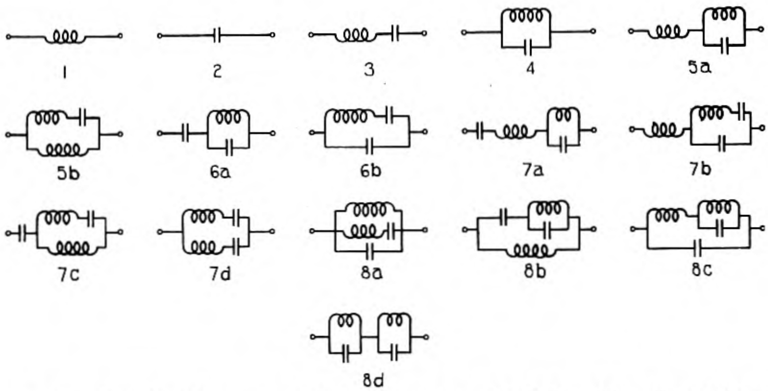


Fig. 5—Two-Terminal Reactance Meshes Containing Not More Than Four Elements

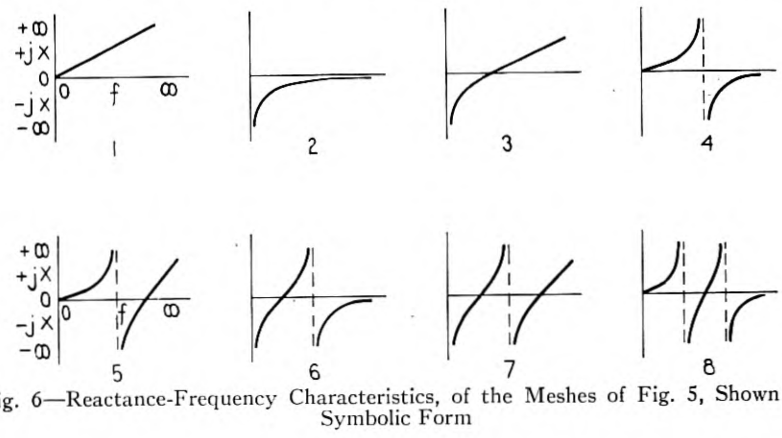


Fig. 6—Reactance-Frequency Characteristics, of the Meshes of Fig. 5, Shown in Symbolic Form

reactance characteristics are represented. Reactance characteristics Nos. 1 and 2 of Fig. 6 are reciprocal in nature, that is, their product is a constant, independent of frequency. Reactance characteristics Nos. 3 and 4 are similarly related if the frequencies of resonance and anti-resonance coincide. Similar relations exist between characteristics Nos. 5 and 6, and between characteristics Nos. 7 and 8. Two forms of reactance mesh in Fig. 5 (Nos. 5a and 5b) give the same reactance characteristic (No. 5 of Fig. 6) and are, therefore, by proper design, electrically equivalent. Characteristic No. 6 of Fig. 6 also corresponds to two reactance meshes of Fig. 5 (Nos. 6a and 6b) and the latter may, therefore, be considered equivalent. Likewise, reactance meshes 7a, 7b, 7c and 7d of Fig. 5 give characteristic No. 7 of Fig. 6 and are therefore potentially equivalent; also reactance

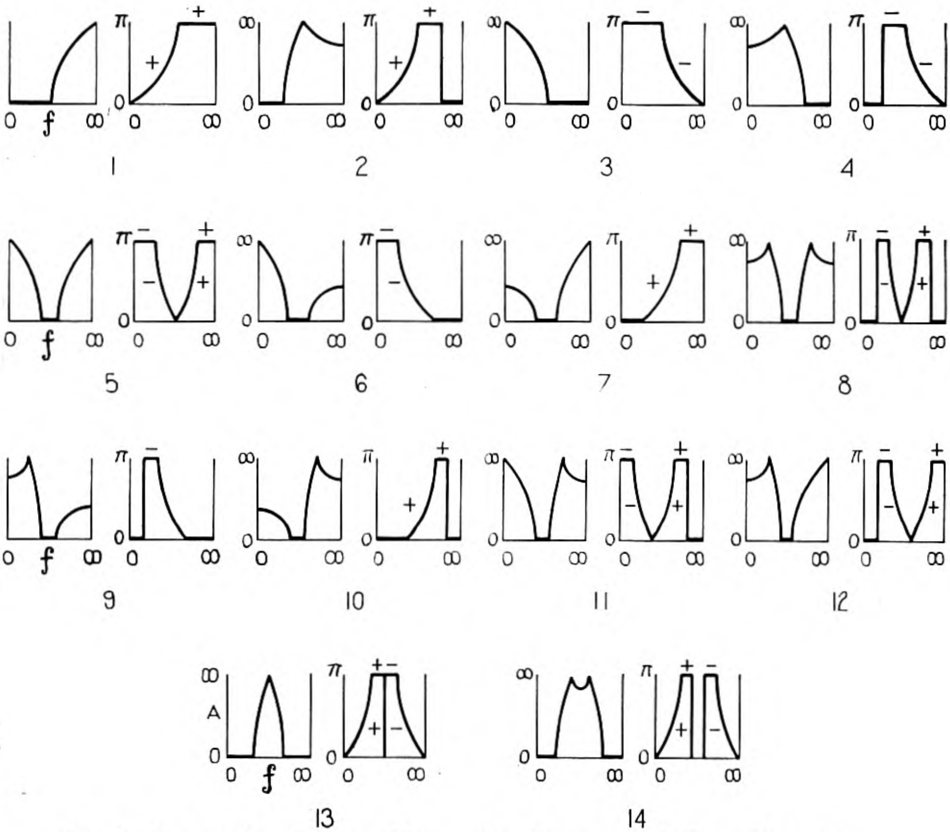


Fig. 7—Propagation Constant (Attenuation Constant and Phase Constant) Characteristics, Shown in Symbolic Form

meshes Nos. 8a, 8b, 8c and 8d of Fig. 5 are represented by reactance characteristic No. 8 of Fig. 6 and, consequently, may also be designed to be equivalent. The equivalence of the above reactance meshes has been discussed by Zobel [5] and will be subsequently treated at length. It is to be understood that, for the sake of brevity, in what follows, meshes Nos. 5, 6, 7 and 8 cover, respectively, all forms of the equivalent meshes: 5a and 5b; 6a and 6b; 7a, 7b, 7c and 7d; and

8a, 8b, 8c and 8d. Using these reactance combinations[9] for the series and shunt arms, there are only a relatively small number of types of filter structures. All of these types of filter structures are

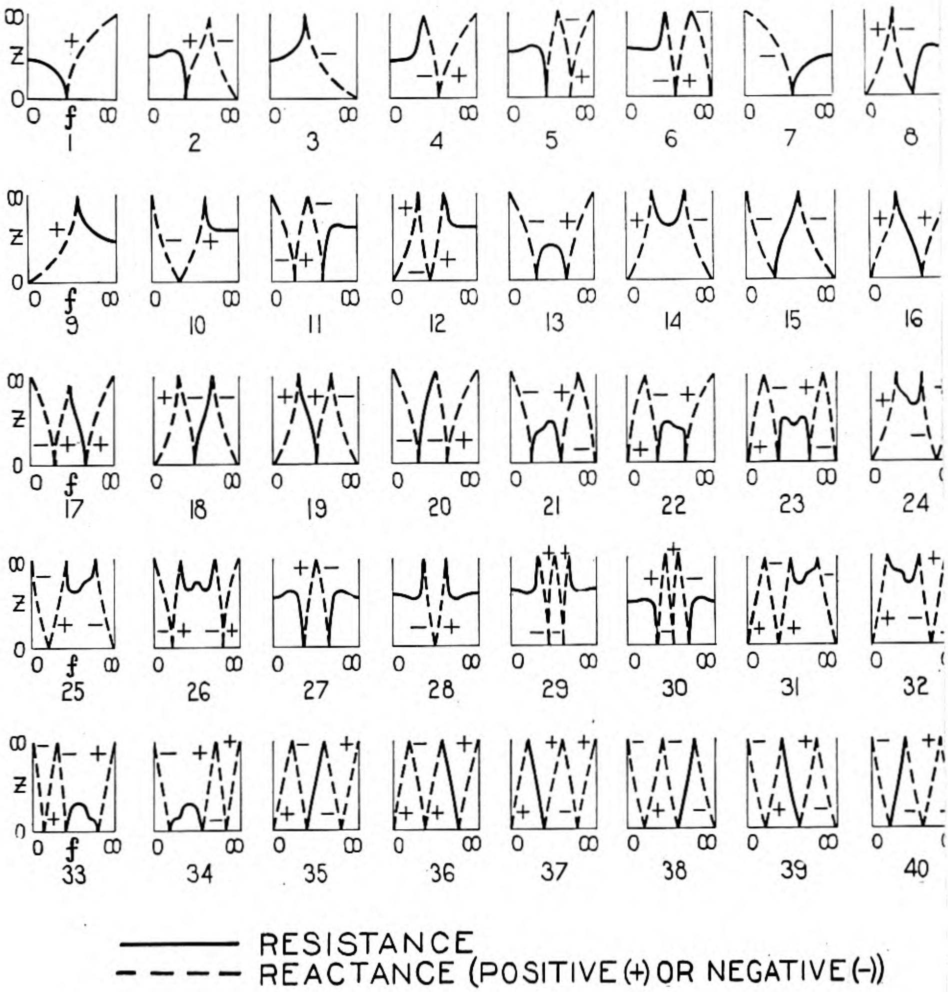


Fig. 8—Mid-Series and Mid-Shunt Image Impedance Characteristics, Shown in Symbolic Form

listed in Table II, and are called *low pass, high pass*, and *band pass* filters (having only one transmission band) and *band elimination*

[9] The general method of deriving the attenuation and phase characteristics of a section from the reactance-frequency characteristics of its series and shunt arms is discussed by Zobel in Bibliography 13.

Tabulation of the Propagation and Impedance Characteristics of Series-Shunt Wave Filter Sections Which can be Formed by the Reactance Meshes Shown in Fig. 5

SERIES ARM

SHUNT ARM	1	2	3	4	5	6	7	8
1	No Pass Band	3–7–9	6–13–16	4–8–9	9–22–16	Band-and-High Pass	Double Band Pass	Band-and-High Pass
2	1–1–3	No Pass Band	7–13–15	2–2–3	Low-and-Band Pass	10–21–15	Double Band Pass	Low-and-Band Pass
3	2–1–4	4–7–11	9–13–17 10–13–20	14–27–28	2–5–4	4–11–10	9–33–17 10–34–20	14–30–28
4	7–16–14	6–15–14	5–13–14	9–18–14 10–19–14	12–22–14	11–21–14	Triple Band Pass	8–23–14
5	10–16–24	Band-and-High Pass	11–13–24	4–8–12	9–22–36 9–35–24 10–22–35 10–37–24	Double Band-and-High Pass	More Than Six Elements	More Than Six Elements
6	Low-and-Band Pass	9–15–25	12–13–25	2–2–6	Low-and-Double Band Pass	9–21–39 9–38–25 10–21–40 10–39–25	More Than Six Elements	More Than Six Elements
7	Low-and-Band Pass	Band-and-High Pass	8–13–26	14–27–29	More Than Six Elements	More Than Six Elements	More Than Six Elements	More Than Six Elements
8	Double Band Pass	Double Band Pass	Triple Band Pass	9–18–31 10–19–32	More Than Six Elements	More Than Six Elements	More Than Six Elements	More Than Six Elements

filters (having two pass bands and only one attenuation band). Their attenuation constant and phase constant characteristics, with respect to frequency, are shown symbolically in Fig. 7. The mid-series and mid-shunt image impedance characteristics with respect to frequency are shown in Fig. 8. In Table II, the figure at the head of each column indicates the reactance mesh in Fig. 5 which is used for Z_1 (series impedance) and the figure at the left of each row indicates the mesh in Fig. 5 which is used for Z_2 (shunt impedance). The figures in the squares of the table denote, reading from left to right, the propagation characteristics (attenuation and phase), the mid-series image impedance, and the mid-shunt image impedance, respectively, as shown in Figs. 7 and 8.

For example, the filter corresponding to the third column and to the fourth row $(3-4)$ has a series arm composed of an inductance in series with a capacity as indicated by mesh 3 of Fig. 5, and has a shunt arm composed of an inductance in parallel with a capacity, as designated by mesh 4 of Fig. 5. The attenuation constant and phase constant characteristics of this filter are shown symbolically by diagram 5 of Fig. 7, while the mid-series and mid-shunt image impedances are indicated, respectively, by diagrams 13 and 14 of Fig. 8. The symbolic nature of the diagrams lies in the fact that the abscissae of each diagram cover the frequency range from zero to infinity, and the ordinates of Figs. 7 and 8 cover the attenuation constant and the impedances from zero to infinity. For example, the structure cited has an attenuation constant characteristic (diagram 5 of Fig. 7) composed of a transmission band lying between two attenuation bands, the attenuation constant being infinite in one of them at zero frequency, and in the other, at infinite frequency. The phase constant of this structure is $-\pi$ radians in the lower of the two attenuation bands, increases from $-\pi$ to $+\pi$ radians in the transmission band (passing through zero), and is $+\pi$ radians throughout the upper of the two attenuation bands. The mid-series image impedance (diagram 13 of Fig. 8) is a negative reactance in the lower of the two transmission bands, decreasing from infinity, at zero frequency, to zero at the lower cut-off frequency, is a pure resistance throughout the transmission band, and is a positive reactance, increasing from zero to infinity, in the upper of the two attenuation bands. The mid-shunt image impedance characteristic (diagram 14 of Fig. 8) is reciprocal in nature, for this structure, to the mid-series image impedance characteristic. This type of filter also possesses, in the general case, a double band pass attenuation characteristic and corresponding phase and impedance characteristics. A discussion of such characteristics is outside the scope of this paper even though many of the structures listed in Table II will show, if completely analyzed, multi-band characteristics. Where no specific characteristics are listed in Table II, no low pass, high pass, single band pass, or single band elimination characteristics are obtainable with a filter section limited to six different reactance elements.

In Table II, a large number of the structures have identically the same types of attenuation constant and phase constant characteristics. For example, six of the seven low pass filter sections have attenuation constant and phase constant characteristic No. 2 of Fig. 7. Likewise, six of the high pass structures have attenuation constant and phase constant characteristic No. 4. Also, in Table II, band pass groups are to be found having respectively, the following propagation characteristics common to each group: 6, 7, 8, 9, 10, 11 and 12. Finally, ten of the eleven band elimination structures listed have propagation constant characteristic No. 14.

Although six of the seven low pass wave filters have the same attenuation constant and phase constant characteristics, the various image impedance characteristics differentiate the structures among themselves. Similar differentiations exist in the high pass, band pass, and band elimination groups of structures. In each of the four types of filter sections however, all of those structures having the same series reactance meshes (that is, having the same series configuration of reactance elements) may be designed to have the same mid-series image impedance characteristic and, similarly, all of those structures within each type having the same shunt reactance meshes, or configuration of elements, may be designed to have the same mid-shunt image impedance characteristic.

In view of the fact that some of the structures listed in Table II have the same attenuation and phase constants but have different impedance characteristics, the question arises as to the relative virtues of the latter. Furthermore, since certain of the structures have the same mid-series or mid-shunt image impedances but have different propagation characteristics, it is possible to join together such structures and obtain a composite structure which has no internal reflection losses, that is, one whose total transfer constant is the sum of the various transfer constants of the individual sections. In order to minimize reflection and interaction losses in the transmission range, it is generally desirable to use, at the terminals of the filter, sections whose image impedances closely simulate those of the terminal impedances to which the filter is connected. The choice presented by filter structures having different impedance characteristics but the same propagation characteristic is, therefore, of advantage. In the attenuation range this is also true where impedance conditions are imposed at the terminals of the filter.

One class of structures which possess desirable image impedances and whose characteristics are readily determined from simpler structures is the so-called derived m-type.[5] The simplest forms of derived

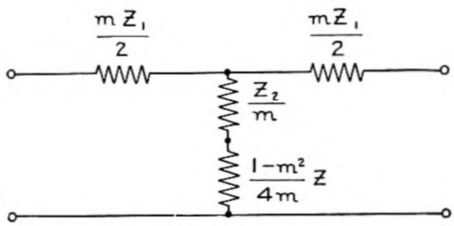


Fig. 9—Mid-Series Equivalent m-Type of Section

structures are shown in Figs. 9 and 10. The structure of Fig. 9 has the same mid-series image impedance as that shown in Fig. 2 and the value of this impedance is given by equation (6). The structure of Fig. 10 has the same mid-shunt image impedance as the π structure shown in Fig. 4 and the value of this impedance is given by

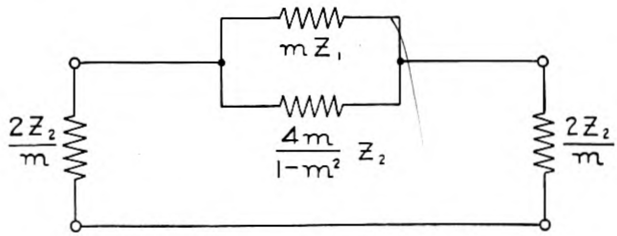


Fig. 10—Mid-Shunt Equivalent m-Type of Section

equation (7). On account of this identity of the respective mid-series and the mid-shunt image impedances in the two cases, the structures shown in Figs. 9 and 10 are called, respectively, the *mid-series equivalent derived m*-type and the *mid-shunt equivalent derived m*-type. The T and π structures of Figs. 2 and 4 are called, respectively, the *prototypes* of the derived m-structures of Figs. 9 and 10. In a series-shunt filter composed of sections of the m-type of Fig. 9 or Fig. 10, the ratio $(Z_1/4Z_2)_m$ of the series impedance to four times the shunt impedance is

$$\left(\frac{Z_1}{4Z_2}\right)_m = \frac{m^2\left(\frac{Z_1}{4Z_2}\right)}{1+(1-m^2)\left(\frac{Z_1}{4Z_2}\right)}. \tag{21}$$

From this expression, when $Z_1/4Z_2$ of the prototype is 0 or -1, the corresponding value of $(Z_1/4Z_2)_m$ for the derived m-type is also 0 or -1. Hence, the derived type has the same cut-off frequencies and

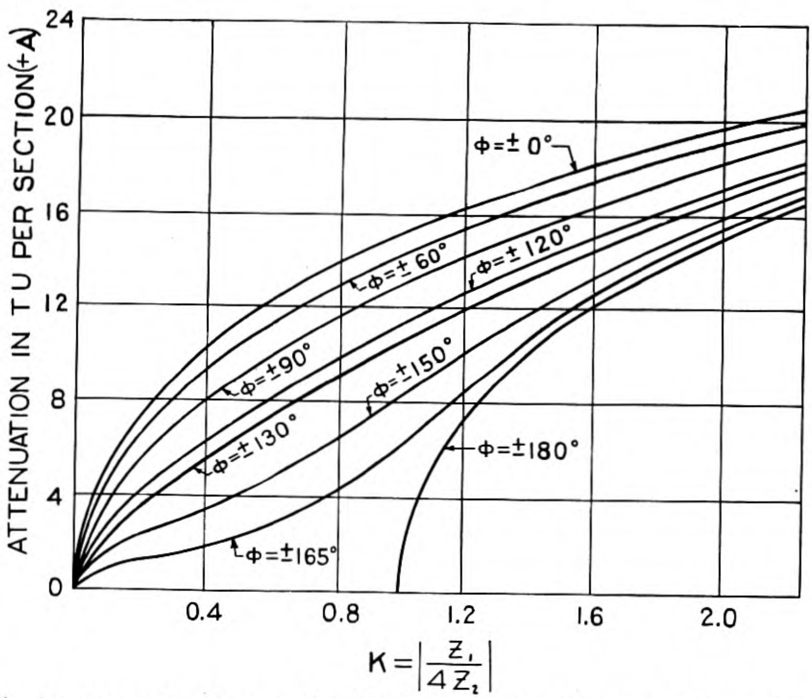


Fig. 11—Attenuation Constant (in TU) of a Filter Section Expressed in Terms of the Ratio of Its Series Impedance to Four Times Its Shunt Impedance (i.e., $Z_1/4Z_2 \equiv K\underline{/\Phi}$)

therefore the same transmission and attenuation regions as its prototype.

Impedance and Propagation Characteristics of Dissipative Filters. It has been pointed out, in the case of non-dissipative structures, that the ratio $Z_1/4Z_2$ is either a positive or a negative numeric. If there is dissipation in the filter structure, that is, if the resistance associated with the reactance elements cannot be neglected, then the ratio

$Z_1/4Z_2$ will not, in general, be a numeric but a vector. However, the general formula (8), still holds true with dissipation. For determining the attenuation constant and phase constant of a dissipative structure it is convenient to use two formulae which may be derived from (8). These formulae are

$$A = \cosh^{-1}\left(K + \sqrt{(K-1)^2 + 4K\cos^2\frac{\phi}{2}}\right), \tag{22}$$

$$B = \cos^{-1}(-K + \sqrt{K^2 + 2K\cos\phi + 1}), \tag{23}$$

where

$$\frac{Z_1}{4Z_2} \equiv \left|\frac{Z_1}{4Z_2}\right| \underline{/\pm\phi} \equiv K\underline{/\pm\phi}.$$

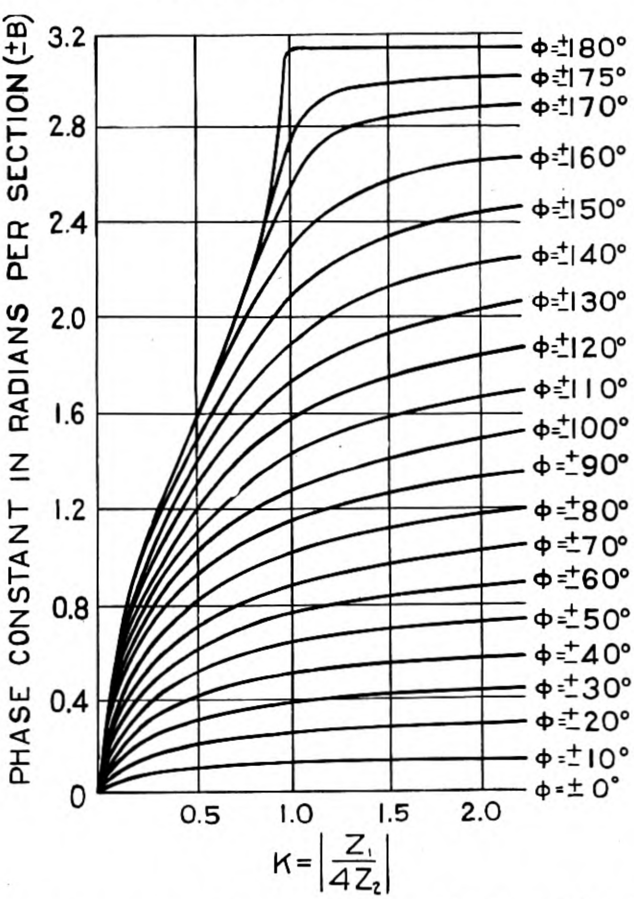


Fig. 12—Phase Constant of a Filter Section Expressed in Terms of the Ratio of Its Series Impedance to Four Times Its Shunt Impedance (i.e., $Z_1/4Z_2 \equiv K_2/\underline{\Phi}$)

Formulae (22) and (23) are expressed in napiers and circular radians, respectively. They are represented in TU and in radians by families of curves such as are shown in Figs. 11 and 12.

A convenient ratio which expresses the dissipation in any reactance element is the absolute ratio, d, of its effective resistance to its reactance. In the case of a coil, $d=R/L\omega$ while in the case of a condenser $d=RC\omega$. The reciprocal ratio $Q\equiv\frac{1}{d}=\frac{L\omega}{R}=\frac{1}{RC\omega}$ has also been widely used as a measure of dissipation in reactance elements. The ratio d or Q will not, in general, be constant over a wide frequency

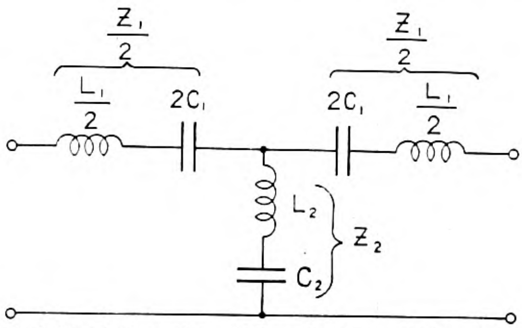


Fig. 13—Typical Band Pass Wave Filter Section (Mid-Series Termination)

range. If the value is known at an important frequency in the transmission range, it may ordinarily be regarded to hold for the rest of the transmission range. The effect of dissipation on the attenuation constant is most important in the transmission band, where the attenuation constant would be zero if there were no dissipation. Its effect is most pronounced in the neighborhood of the cut-off frequencies where the transmission bands merge into attenuation bands.

In the attenuation bands, the general effect of dissipation is negligible. It largely controls, however, the value of the attenuation constant at those frequencies at which infinite attenuation would occur if there were no dissipation. The effect of dissipation upon the phase constant is most pronounced in the neighborhood of the cut-off frequencies where resistance rounds off the abrupt changes in phase which would otherwise occur (see Fig. 12).

Characteristics of a Typical Filter. In order to illustrate specifically the principles employed in filter design, consider as an example the band pass structure 3−3 of Table II. This structure is illustrated in Fig. 13. It will be assumed that the dissipation in the coils cannot be neglected, but that the dissipation in the condensers is of negligible magnitude. If R_1 and R_2 are the effective resistances of the inductance elements L_1 and L_2, respectively, the series impedance, Z_1, of a series-shunt recurrent structure composed of sections of the type shown in Fig. 13 is

$$Z_1=R_1+j\left(\omega L_1-\frac{1}{\omega C_1}\right). \tag{24}$$

The impedance of the shunt arm is

$$Z_2=R_2+j\left(\omega L_2-\frac{1}{\omega C_2}\right). \tag{25}$$

In substituting for R_1 its value $L_1\omega d$ and for R_2 its value $L_2\omega d$, the ratio $Z_1/4Z_2$ becomes

$$\frac{Z_1}{4Z_2}=\frac{L_1}{4L_2}\,\frac{1-jd-\dfrac{1}{\omega^2 L_1 C_1}}{1-jd-\dfrac{1}{\omega^2 L_2 C_2}}. \tag{26}$$

Assuming d to be zero, the ratio $Z_1/4Z_2$ is

$$\frac{Z_1}{4Z_2}=\frac{C_2\,(\omega^2 L_1 C_1-1)}{4C_1(\omega^2 L_2 C_2-1)}. \tag{27}$$

Referring to Table II, the structure shown in Fig. 13 has two distinct attenuation and phase characteristics. These are, respectively, characteristics Nos. 9 and 10 of Fig. 7. These two sets of characteristics arise from the fact that the shunt arm may be resonant at a frequency less than, or greater than, the resonant frequency of the series arm. The two attenuation characteristics are inverse with respect to frequency. We shall, therefore, discuss only one of the two cases, namely, that in which the shunt arm resonates at a frequency greater than the resonant frequency of the series arm (that is, L_1C_1 is greater than L_2C_2). The frequency at which the shunt arm is resonant will be designated as f_∞, due to the fact that in a non-dissipative filter the attenuation constant is infinite at this point. In other words,

$$f_\infty=\frac{1}{2\pi\sqrt{L_2C_2}}. \tag{28}$$

It is evident that the frequency at which Z_1 is resonant is a cut-off frequency since Z_1, and therefore $Z_1/4Z_2$, is zero at this point. An inspection of graphical curves [10] drawn for Z_1 and $4Z_2$, under the above

[10] For an illustration of the construction of such curves see Bibliography 12, Fig. 7, also Bibliography 13, Fig. 2.

conditions, will show that this is the lower of the two cut-off frequencies (f_1), that is

$$f_1=\frac{1}{2\pi\sqrt{L_1C_1}}. \tag{29}$$

By equating $Z_1/4Z_2$ to -1 in equation (27) the upper cut-off frequency (f_2) is found to be

$$f_2=\frac{1}{2\pi}\sqrt{\frac{C_2+4C_1}{C_1C_2(L_1+4L_2)}}. \tag{30}$$

For these explicit relations for f_1, f_2 and f_∞, equation (26) may be rewritten

$$\frac{Z_1}{4Z_2}=\left(\frac{f_1}{f_\infty}\right)^2\frac{\left[\left(\frac{f_\infty}{f_2}\right)^2-1\right]\left[(1-jd)\left(\frac{f}{f_1}\right)^2-1\right]}{\left[1-\left(\frac{f_1}{f_2}\right)^2\right]\left[(1-jd)\left(\frac{f}{f_\infty}\right)^2-1\right]}. \tag{31}$$

When d is zero this equation becomes, for the non-dissipative case

$$\frac{Z_1}{4Z_2}=\frac{\left[1-\left(\frac{f_\infty}{f_2}\right)^2\right]\left[1-\left(\frac{f_1}{f}\right)^2\right]}{\left[1-\left(\frac{f_1}{f_2}\right)^2\right]\left[\left(\frac{f_\infty}{f}\right)^2-1\right]}. \tag{32}$$

From the preceding formulae and from the curves shown in Figs. 11 and 12, it is possible to read directly the attenuation constant and the phase constant for the structure shown in Fig. 13, at any frequency, provided the values of f_1, f_2 and f_∞ are known. The formulae for the dissipative case are of use mainly throughout the transmission bands and near the frequency f_∞. Elsewhere, the formulae for $Z_1/4Z_2$ for the non-dissipative structure may be employed without undue error. The preceding formulae have been derived in a direct manner, but may be obtained more simply by considering the structure of Fig. 13 to be a derived form of the structure $3-2$ in Table II.

In order to minimize reflection loss effects, it is, as a rule, desirable to terminate a filter in an impedance equal to the image impedance of the filter at the mid-frequency,[11] (f_m) or at some other important frequency. From equation (6) and the values of Z_1 and Z_2, the mid-series image impedance (Z_o), at the mid-frequency in the non-dissipative case is

$$Z_o=\frac{1}{2}\left[\sqrt{\frac{L_1}{C_1}+4\frac{L_1}{C_2}}-\sqrt{\frac{L_1}{C_1}+4\frac{L_2}{C_1}}\right]. \tag{33}$$

[11] Defined as the geometric mean of the two cut-off frequencies f_1 and f_2; or $f_m \equiv \sqrt{f_1 f_2}$.

From formulae (6), (29), (30), and (33) the mid-series image impedance at any frequency is

$$Z_I = Z_o \sqrt{1 - \frac{\left(\frac{f}{f_m} - \frac{f_m}{f}\right)^2}{\left(\frac{f_2}{f_m} - \frac{f_1}{f_m}\right)^2}}. \tag{34}$$

An inspection of formula (34) indicates that the mid-series image impedance is symmetrical with respect to the mid-frequency, f_m.

In a similar way, the mid-shunt image impedance (Z_o') at the mid-frequency is

$$Z_o' = \sqrt{\frac{4L_1}{C_2\left(\frac{C_2}{C_1}+4\right)}} - \sqrt{\frac{4L_2}{C_1\left(\frac{L_1}{L_2}+4\right)}} \tag{35}$$

and the mid-shunt impedance, (Z_I'), at any frequency is

$$Z_I' = Z_o' \frac{1-\left(\frac{f}{f_\infty}\right)^2}{1-\left(\frac{f_m}{f_\infty}\right)^2} \sqrt{\frac{\frac{f_2}{f_1} - \left(\frac{f_m}{f}\right)^2}{\frac{f_2}{f_1} - \left(\frac{f}{f_m}\right)^2}} \tag{36}$$

It will be noted, that *if the values of the inductances and resistances of a filter are multiplied by any factor and if all the values of the capacities are divided by the same factor, the transmission loss-frequency characteristic is not changed*[12] *(neither are the cut-off frequencies, nor the frequencies of infinite attenuation) but the image impedances are multiplied by this factor.*

From the preceding formulae, explicit expressions may be derived for the values of L_1, C_1, L_2, and C_2. These expressions, which are given by Zobel,[5] in a slightly different form, are as follows:

$$L_1 = \frac{Z_o m}{\pi(f_2 - f_1)}, \tag{37}$$

$$C_1 = \frac{f_2 - f_1}{4\pi f_1^2 Z_o m}, \tag{38}$$

$$L_2 = \frac{Z_o}{\pi(f_2 - f_1)} \frac{1-m^2}{4m}, \tag{39}$$

$$C_2 = \frac{(f_2 - f_1)m}{\pi Z_o (f_2^2 - f_1^2 m^2)}, \tag{40}$$

[12] Since the value of the transfer factor, $\epsilon^{-\theta}$, is dependent simply upon the ratio $Z_1/4Z_2$, it is evident from equation (10) that *the transmission loss caused by the insertion of any network in a circuit is dependent simply upon impedance ratios.* Consequently, the above theorem is quite general and applies not only to filters but to any passive network.

where

$$m=\sqrt{1-\frac{\left(\frac{f_2}{f_1}\right)^2-1}{\left(\frac{f_\infty}{f_1}\right)^2-1}}. \tag{41}$$

As a numerical example of the determination of the constants of a filter section of the type under consideration, assume that the lower cut-off frequency, f_1, is 20,000 cycles, and that the upper cut-off frequency, f_2, is 25,000 cycles and that the frequency of infinite attenuation, f_∞, is 30,000 cycles. Assume, furthermore, that the value of the mid-series image impedance, Z_o, at the mid-frequency is 600 ohms. Then from formula (41), $m=.742$; hence from (37), $L_1=.0284$ henry; from (38), $C_1=.00224\times10^{-6}$ farad; from (39) $L_2=.00577$ henry and from (40) $C_2=.00486\times10^{-6}$ farad. Assuming $d=.01$, the value of $Z_1/4Z_2$ as given by formula (31) at f_m (22,360 cycles) is found to be $.305\underline{/176°.4}$. Referring to formula (22), in which $K=.305$ and $\phi=176°.4$, or to the curves of Fig. 11, this value of $Z_1/4Z_2$ corresponds approximately to .041 napiers or .36 *TU*. Similarly, from equation (23), or from the curves of Fig. 12, this value of $Z_1/4Z_2$ gives 1.15 radians, or 67°, for the phase constant. At zero frequency, the value of $Z_1/4Z_2$ is, from equation (31), $.542\underline{/0°}$, which corresponds to 1.36 napiers or to 11.8 *TU*. Likewise, at infinite frequency, the value of $Z_1/4Z_2$ is $1.23\underline{/0°}$, which corresponds to an attenuation loss of 1.97 napiers or to 16.6 *TU*. From the curves of Fig. 12, the phase constant is zero both at zero and at infinite frequency.

Composite Wave Filters. It has previously been pointed out that certain groups of the structures listed in Table II have the same mid-series or mid-shunt image impedance characteristics but that the various structures in such a group may have different attenuation and phase constant characteristics.

If a filter is composed of any number of symmetrical or dissymmetrical sections, so joined together that the image impedances at the junction points of the sections are identical, the attenuation and phase constant characteristics of the composite structure so formed, are equal to the sum of the respective characteristics of the individual sections. Furthermore, the image impedances of the composite filter will be determined by the image impedances of the accessible ends of the terminating sections. The desirability of forming such composite filters arises from the fact that a better disposition of attenuation and phase can be obtained by employing, in one composite structure, a number of different types of the characteristics shown in Fig. 7.

The dissymmetrical networks ordinarily employed in composite structures are usually L type networks each of which may be regarded as one-half the corresponding symmetrical T or π network. Generalized forms of such networks are shown in Figs. 14A, B, and C. By joining two of these half-sections, such as are shown in Figs. 14B

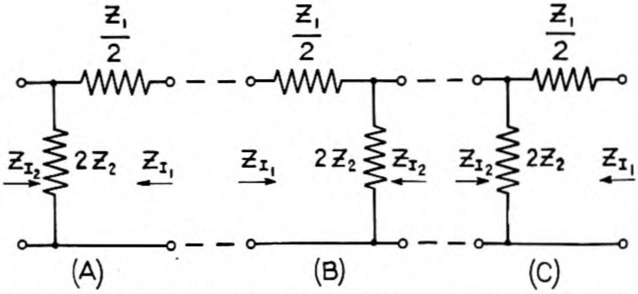


Fig. 14—Generalized Series-Shunt Structure Divided Into Successive Half-Sections (L-Type)

and C, we may form the full T section shown in Fig. 2. Similarly, by joining the two half-sections illustrated in Figs. 14A and B, the full π section of Fig. 4 results. The transfer constant, $\Theta_{1/2}$, of a half-section, such as is shown in Figs. 14A, B, or C, is one-half the transfer constant of the corresponding full section, that is,

$$\Theta_{1/2} = \frac{\Theta}{2} = \sinh^{-1}\sqrt{\frac{Z_1}{4Z_2}}. \tag{42}$$

Hence, the *attenuation constant and phase constant of a half-section* are, respectively, *one-half the attenuation constant and phase constant of a full section*. An important relationship between the half-section and the full section, which makes it convenient to use half-sections in composite wave filter structures, is that the image impedances, Z_{I_1} and Z_{I_2}, of any half-section are equal respectively to the mid-series and the mid-shunt image impedances of the corresponding full sections.

A typical example of the method of forming a composite low pass wave filter is given in Fig. 15, where three half-sections of different types and one full section are combined into a composite filter. The designations below the diagrams in Fig. 15A refer to the number of full sections and to the ratio f_∞/f_c. In a practical filter, the various shunt condensers and series coils are combined as illustrated in Fig. 15B.

The composite nature of the attenuation characteristic of the filter of Fig. 15B is illustrated in Fig. 16, on a non-dissipative basis. In

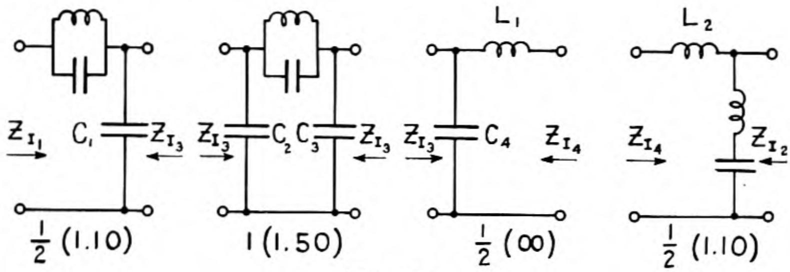


Fig. 15 A

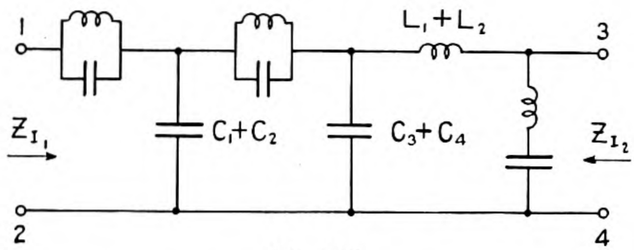


Fig. 15 B

Typical (Non-Dissipative) Composite Low Pass Wave Filter and Its Component Sections and Half-Sections

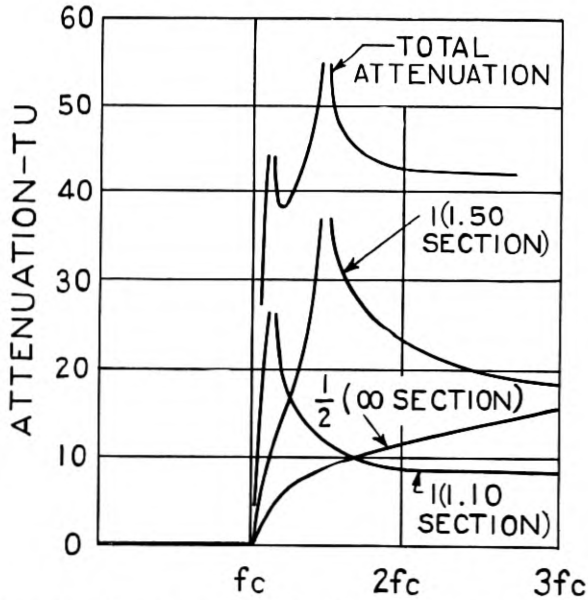


Fig. 16—Attenuation Characteristic of the Composite Low Pass Wave Filter of Fig. 15

Fig. 15B, the image impedance, Z_{I_1}, at the $1-2$ terminals has characteristic No. 2 of Fig. 8, while the image impedance, Z_{I_2}, at the $3-4$ terminals has characteristic No. 4 of Fig. 8.

Electrically Equivalent Networks. Reference has been made to the fact that *any passive network having one pair of input terminals and one pair of output terminals may be adequately represented, at any frequency, by an equivalent T or π network.* In general, this representation is a mathematical one and the arms of the T or π network cannot be represented, at all frequencies, by *physically realizable* impedances.

Furthermore, *any concealed network, containing no impressed electromotive forces, and having N accessible terminals is always capable of mathematical representation, at a single frequency, by a network having not more than $N\ (N-1)/2$ impedances, which impedances are determinable from the voltage and current conditions at the accessible terminals.* For networks having three or more terminals, this arbitrary mesh of impedances may possess a number of variant configurations. It is also true that the equivalence of the arbitrary mesh to the concealed network holds, at any single frequency, for any and all sets of external or terminal conditions, and that the magnitudes of the impedances of the arbitrary mesh are determinable, at will, on the assumption of the most convenient set of terminal conditions for each individual case. Familiar instances are the impedance equations derivable under various short-circuit and open-circuit conditions.

In specific cases, which are of particular interest, *one network may be shown to be capable of representation, as far as external circuit conditions are concerned, by another network which is physically realizable, and the latter may be substituted for the former, indiscriminately, in any circuit without consequent alteration, at any frequency, in the circuit conditions external to the interchanged networks.*

Equivalent meshes having two accessible terminals and employing respectively, three or four impedances in each mesh have been discussed by O. J. Zobel.[13] In filter design, two-terminal meshes are of importance only in those cases where the impedances are essentially reactances. Figs. 17A, B, C and D illustrate the physical configurations which reactance meshes employing not more than four elements may take. We are not generally interested in meshes having more than four elements for practical reasons which have previously been discussed. *Whenever any of the reactance meshes shown in Fig. 17 occur, we may, with proper design, substitute for it an equivalent mesh*

[13] See Appendix III of Bibliography 13.

of the associated type or types. Rigorous equivalence exists, *even with dissipation,* when the ratio of resistance to reactance, (d), is the same for all coils and the ratio of resistance to reactance (d') is the same for all condensers.

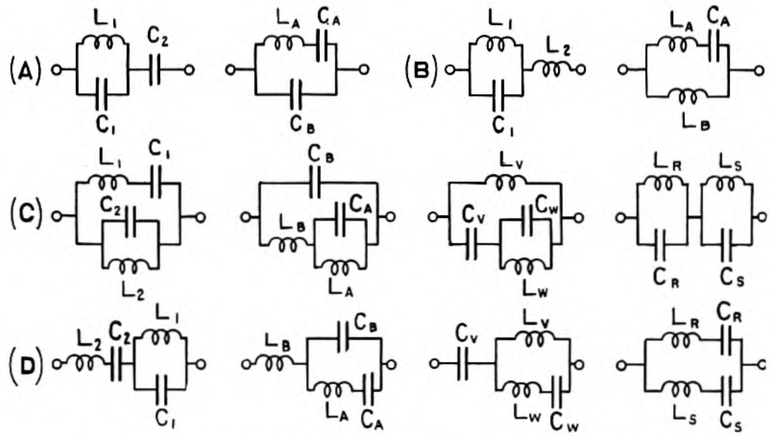


Fig. 17—Groups of Equivalent Two-Terminal Reactance Meshes

The relations which the equivalent meshes of Fig. 17 must observe are as follows:

17A

$$C_2=C_A+C_B,\ C_1=\frac{C_B}{C_A}(C_A+C_B),\ L_1=\frac{L_A}{\left(1+\frac{C_B}{C_A}\right)^2}, \tag{43}$$

$$C_A=\frac{C_2^2}{C_1+C_2},\ C_B=\frac{C_1\ C_2}{C_1+C_2},\ L_A=L_1\left(1+\frac{C_1}{C_2}\right)^2, \tag{44}$$

17B

$$L_1=\frac{L_B^2}{L_A+L_B},\ C_1=C_A\left(1+\frac{L_A}{L_B}\right)^2,\ L_2=\frac{L_A\ L_B}{L_A+L_B}, \tag{45}$$

$$L_A=\frac{L_2}{L_1}(L_1+L_2),\ C_A=\frac{C_1}{\left(1+\frac{L_2}{L_1}\right)^2},\ L_B=L_1+L_2, \tag{46}$$

17C

$$L_1=\frac{L_B}{L_A}(L_A+L_B)=L_W\left(1+\frac{C_W}{C_V}\right)^2 =\frac{L_RL_S(L_R+L_S)(C_R+C_S)^2}{(L_RC_R-L_SC_S)^2} \tag{47}$$

$$L_2=L_A+L_B=L_V=L_R+L_S, \tag{48}$$

17C

$$C_1 = \frac{C_A}{\left(1+\frac{L_B}{L_A}\right)^2} = \frac{C_V^2}{C_V+C_W} = \frac{(L_R C_R - L_S C_S)^2}{(L_R+L_S)^2\,(C_R+C_S)}, \tag{49}$$

$$C_2 = C_B = \frac{C_V C_W}{C_V+C_W} = \frac{C_R C_S}{C_R+C_S}, \tag{50}$$

$$L_A = \frac{L_2^2}{L_1+L_2},\ L_B = \frac{L_1 L_2}{L_1+L_2},\ C_B = C_2, \tag{51}$$

$$C_A = C_1\left(1+\frac{L_1}{L_2}\right)^2,\ L_W = \frac{L_1}{\left(1+\frac{C_2}{C_1}\right)^2}, \tag{52}$$

$$L_V = L_2,\ C_V = C_1 + C_2,\ C_W = \frac{C_2}{C_1}(C_1+C_2), \tag{53}$$

$$C_S = \frac{K+\sqrt{K^2-4L_2^2 C_1 C_2 K}}{2L_2^2 C_1}$$
$$\text{where } K = (L_1C_1+L_2C_1+L_2C_2)^2 - 4L_1C_1L_2C_2, \tag{54}$$

$$C_R = \frac{C_S C_2}{C_S - C_2},\ L_S = \frac{L_1C_1+L_2C_1+L_2C_2-L_2C_R}{C_S-C_R},\ L_R = L_2 - L_S, \tag{55}$$

17D

$$C_1 = \frac{C_B}{C_A}(C_A+C_B) = C_W\left(1+\frac{L_W}{L_V}\right)^2 = \frac{C_R C_S (C_R+C_S)(L_R+L_S)^2}{(L_R C_R - L_S C_S)^2}, \tag{56}$$

$$C_2 = C_A + C_B = C_V = C_R + C_S, \tag{57}$$

$$L_1 = \frac{L_A}{\left(1+\frac{C_B}{C_A}\right)^2} = \frac{L_V^2}{L_V+L_W} = \frac{(L_R C_R - L_S C_S)^2}{(C_R+C_S)^2 (L_R+L_S)}, \tag{58}$$

$$L_2 = L_B = \frac{L_V L_W}{L_V+L_W} = \frac{L_R L_S}{L_R+L_S}, \tag{59}$$

$$C_A = \frac{C_2^2}{C_1+C_2},\ C_B = \frac{C_1 C_2}{C_1+C_2},\ L_B = L_2, \tag{60}$$

$$L_A = L_1\left(1+\frac{C_1}{C_2}\right)^2,\ C_W = \frac{C_1}{\left(1+\frac{L_2}{L_1}\right)^2}, \tag{61}$$

$$C_V = C_2,\ L_V = L_1 + L_2,\ L_W = \frac{L_2}{L_1}(L_1+L_2), \tag{62}$$

$$L_S = \frac{K+\sqrt{K^2-4L_1L_2C_2^2K}}{2L_1C_2^2}$$
$$\text{where } K = (L_1C_1+L_1C_2+L_2C_2)^2 - 4L_1C_1L_2C_2, \tag{63}$$

$$L_R = \frac{L_S L_2}{L_S - L_2},\ C_S = \frac{L_1C_1+L_1C_2+L_2C_2-L_RC_2}{L_S-L_R},\ C_R = C_2 - C_S. \tag{64}$$

For example, the two meshes in Fig. 17A will be equivalent if

$C_1 = .009$ mf.	$C_2 = .001$ mf.	$L_1 = .001$ h.
$C_B = .0009$ mf.	$C_A = .0001$ mf.	$L_A = .100$ h.

and the two meshes in Fig. 17B will be equivalent if

$L_1 = .002$ h.	$C_1 = .025$ mf.	$L_2 = .008$ h.
$L_A = .040$ h.	$C_A = .001$ mf.	$L_B = .010$ h.

Also, the four meshes of Fig. 17C will be equivalent if

$L_R = .001$ h.	$L_S = .002$ h.	$C_R = .001$ mf.	$C_S = .002$ mf.
$L_1 = .006$ h.	$L_2 = .003$ h.	$C_1 = .000333$ mf.	$C_2 = .000667$ mf.
$L_A = .001$ h.	$L_B = .002$ h.	$C_A = .003$ mf.	$C_B = .000667$ mf.
$L_V = .003$ h.	$L_W = .000667$ h.	$C_V = .001$ mf.	$C_W = .002$ mf.

and the four meshes of Fig. 17D will be equivalent if

$L_R = .001$ h.	$L_S = .001$ h.	$C_R = .001$ mf.	$C_S = .002$ mf.
$L_1 = .0000555$ h.	$L_2 = .0005$ h.	$C_1 = .024$ mf.	$C_2 = .003$ mf.
$L_A = .0045$ h.	$L_B = .0005$ h.	$C_A = .000333$ mf.	$C_B = .00267$ mf.
$L_V = .000555$ h.	$L_W = .005$ h.	$C_V = .003$ mf.	$C_W = .00024$ mf.

It is then evident that the following reactance meshes of Fig. 5 may be designed to be equivalent: 5a and 5b; 6a and 6b; 7a, 7b, 7c, and 7d; and 8a, 8b, 8c, and 8d. Hence, the following filter sections

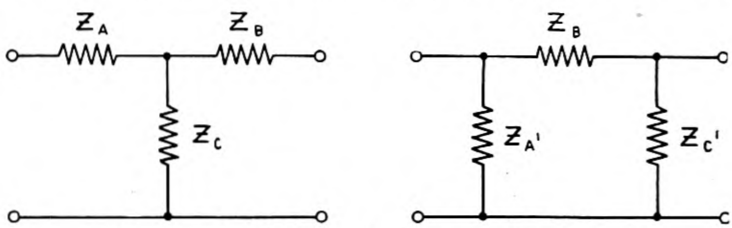


Fig. 18—Equivalent T and π Generalized Networks

referred to in Table II have, for the same impedance and propagation characteristics, a number of variant forms of physical configuration. 4−6, 6−2, 3−5, 6−4, 2−6, 5−3, 4−5, 1−5, 3−6, 5−4, 5−1, 4--8, 5−5, 6−6, 7−3, 6−3, 3−7, 4−7, 8−4 and 8−3.

Of the equivalent meshes having three accessible terminals the most common are the familiar T and π networks. The general relationships which must be observed for the equivalence of T or π networks are due to Kennelly [14] and for their generalized form, as illustrated in Fig. 18, are as follows:

$$Z_A=\frac{Z_A'Z_B'}{Z_A'+Z_B'+Z_C'},\quad Z_B=\frac{Z_B'Z_C'}{Z_A'+Z_B'+Z_C'},\quad Z_C=\frac{Z_A'Z_C'}{Z_A'+Z_B'+Z_C'},\tag{65}$$

$$Z_A'=Z_A+Z_C+\frac{Z_AZ_C}{Z_B},\ Z_B'=Z_A+Z_B+\frac{Z_AZ_B}{Z_C},\ Z'_C=Z_B+Z_C+\frac{Z_BZ_C}{Z_A}.\tag{66}$$

We shall discuss here only two of the principal reactance meshes of the T and π form, namely, those employing solely inductances and

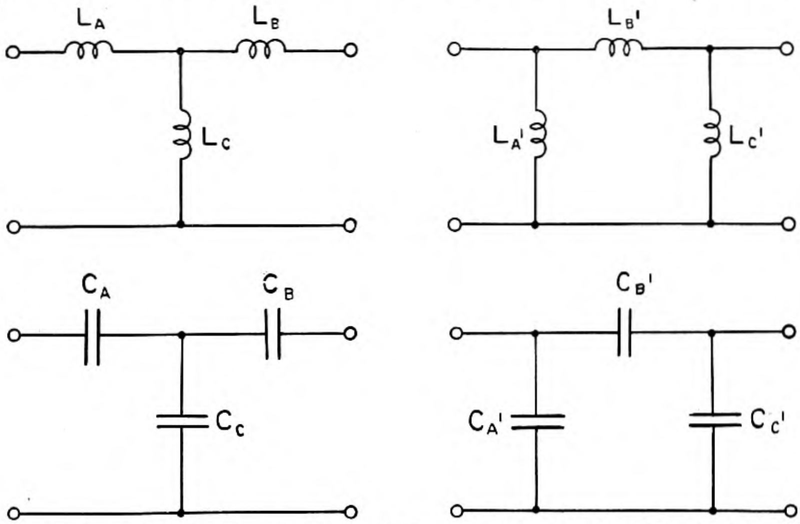


Fig. 19—Equivalent T and π Inductance Networks and Equivalent T and π Capacity Networks

solely capacities. It is to be understood that wherever an inductance or a capacity mesh of any of the following types occurs, its variant network may be substituted for it without change in the electrical characteristics of the circuit excluding those conditions within the mesh or its variant. Fig. 19 illustrates equivalent T and π networks of inductance and capacity.[15] The formulae relating the inductance and capacity meshes of Fig. 19 are as follows:

$$L_A=\frac{L_A'L_B'}{L_A'+L_B'+L_C'},\quad L_B=\frac{L_B'L_C'}{L_A'+L_B'+L_C'},\quad L_C=\frac{L_A'L_C'}{L_A'+L_B'+L_C'},\tag{67}$$

[14] Kennelly, A. E., "The Equivalence of Triangles and Three-Pointed Stars in Conducting Networks," *Electrical World and Engineer*, New York, Vol. XXXIV, No. 12, pp. 413–414, Sept. 16, 1899. Also, "Application of Hyperbolic Functions to Electrical Engineering" (1911) (Appendix E).

[15] These meshes are rigorously equivalent, even when resistance is present if the ratio d is the same for all of the inductances and if the ratio d' is the same for all of the capacities.

$$L_A' = L_A + L_C + \frac{L_A L_C}{L_B},\ L_B' = L_A + L_B + \frac{L_A L_B}{L_C},\ L_C' = L_B + L_C + \frac{L_B L_C}{L_A}, \quad (68)$$

$$C_A' = \frac{C_A C_C}{C_A + C_B + C_C}, \quad C_B' = \frac{C_A C_B}{C_A + C_B + C_C}, \quad C_C' = \frac{C_B C_C}{C_A + C_B + C_C}, \quad (69)$$

$$C_A = C_A' + C_B' + \frac{C_A' C_B'}{C_C'}, \qquad C_B = C_B' + C_C' + \frac{C_B' C_C'}{C_A'},$$

$$C_C = C_A' + C_C' + \frac{C_A' C_C'}{C_B'}. \quad (70)$$

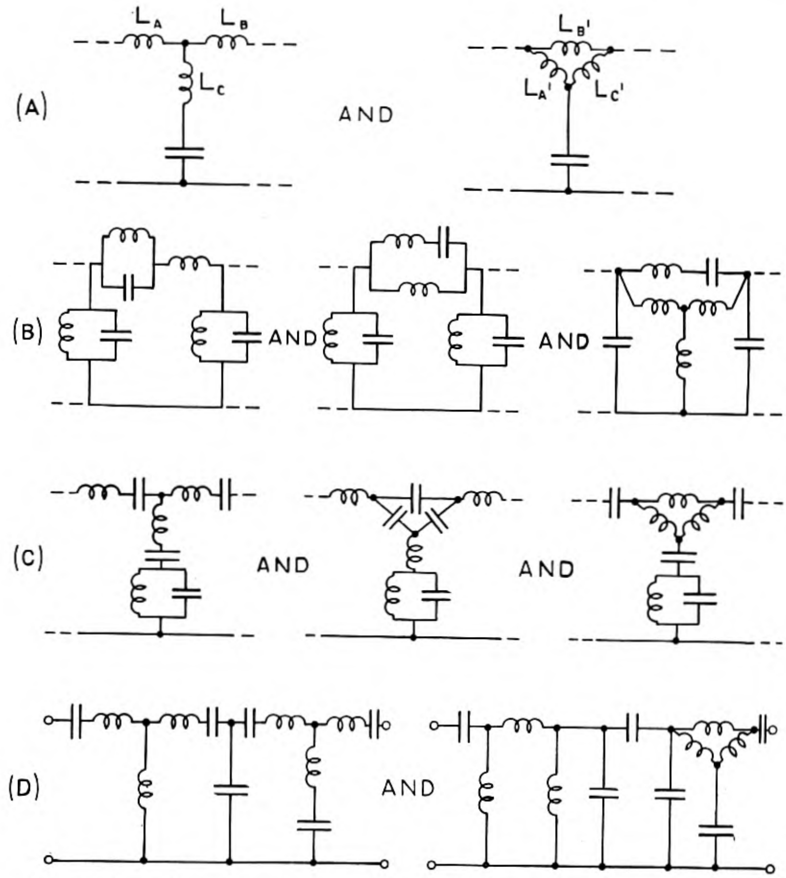


Fig. 20—Typical Examples of Equivalent Filters Involving the Interchange of Three-Terminal Networks of Inductances or of Capacities

A few examples of the variant filter structures which may arise, due to the existence of equivalent three terminal meshes of capacity and inductance, are illustrated in Fig. 20, in which Figs. 20A, B, and C represent either individual sections or portions of composite filters and Fig. 20D represents a composite filter. When equivalent reactance meshes occur entirely within a filter or within a section of a filter, the filter or the section will have the same cut-off frequencies and frequencies of infinite attenuation and the same attenuation, phase, and image impedance characteristics, whichever equivalent

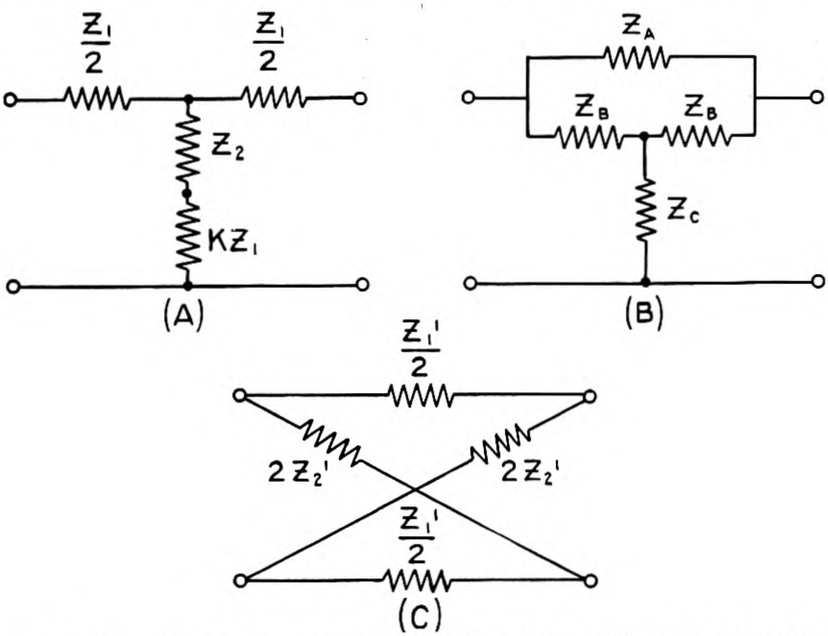


Fig. 21—Generalized Forms of Equivalent Series-Shunt, Bridged-T, and Lattice Type Filter Structures

form of mesh is substituted for an existing mesh. When equivalent meshes are interchanged in either recurrent or composite filters the substitution is generally made after the series-shunt structure is designed and after it has been found that the substitution will effect economies. The three terminal meshes referred to occur, in general, in unbalanced filter structures. For balanced filter circuits, corresponding meshes will be found for each of the equivalent networks by the process of dividing equally the series impedance between the two series lines of the filter.

While the discussion in this paper is based principally on the series-shunt structure there are two other important types of structures which will be mentioned. These are the so-called *lattice* [5] type structure and the *bridged-T* type structure. Typical series-shunt, bridged-T, and lattice type structures are illustrated in Fig. 21A, B and C, respectively. The three circuits shown are electrically equivalent, except for balance between the series arms, if the following relations hold:

$$Z_A = \left(1+\frac{1}{4K}\right)Z_1, \qquad Z_B = (\tfrac{1}{2}+2K)Z_1, \qquad Z_C = Z_2, \tag{71}$$

$$Z_1' = Z_1, \qquad Z_2' = (\tfrac{1}{4}+K)Z_1 + Z_2. \tag{72}$$

In the previous discussion of equivalent networks no reference has been made to networks containing mutual inductance, many of which are of particular interest and importance. These will be now discussed in detail.

PART II

WAVE FILTERS USING MUTUAL INDUCTANCE

Before considering the equivalent meshes which may be formed by the use of mutual inductance between pairs of coils, and the types of wave filters which may be obtained by the use of these equivalent meshes, it will be necessary to define certain general terms.

The *self impedance* between any two terminals of an electrical network is the vector ratio of an applied e.m.f. to the resultant current entering the network when all other accessible terminals are free from external connections.

The *mutual impedance* of any network, having one pair of input terminals and one pair of output terminals, is the *vector ratio* of the e.m.f. produced at the output terminals of the network, on open circuit, to the current flowing into the network at the input terminals. Since mutual impedance is a vector ratio, it may have either of two signs, depending on the assumed directions of the input current and the output voltage. The sign of the mutual impedance is, in general, identified by its effect in increasing or decreasing the vector impedance of the meshes in which it exists. It is usually convenient, in this case, to consider either a simple series or a simple parallel mesh of two self impedances between which the mutual impedance acts. For the purpose of determining the sign of the mutual impedance, we shall confine our discussion to a simple series combination. Consequently, the mutual impedance will be called either *series aiding* or *series opposing*.

When a mutual impedance, Z_M, acts between two self impedances Z_1 and Z_2, (Fig. 22) connected in series in such a way as to *increase vectorially* the impedance of the combination, it is called a *series aiding* losses) the arms of its equivalent T network are composed simply of positive or negative inductances. Of the three inductances involved, at least two of them must be positive while the third may be either positive or negative.

From Fig. 25, it is evident that two windings or coils, together with their mutual impedance, may be represented by an equivalent network which affords a transfer of energy from one winding to the other. This equivalent network may, with limitations, contain positive or negative inductances.

While the two-winding transformer of Fig. 23 has been represented by an equivalent T network in Fig. 26, the equivalent network may alternatively be of π form (Fig. 27) instead of T form, through the

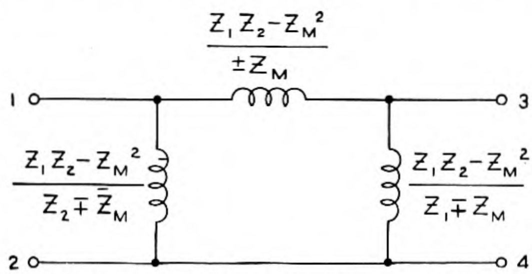


Fig. 27—π Network of Self Impedances Equivalent to the Structure of Fig. 23

general relationships for T or π networks previously stated. When no dissipation exists in the transformer, either equivalent network will have at least two positive inductances while the third inductance may be either positive or negative.

From the principles previously outlined in Part I, for the equivalence of certain electrical meshes and for their substitution for one another in any circuit, it is obvious that when two coils, with mutual impedance between them, exist in a circuit, in the manner shown in Fig. 23, either of the meshes shown in Fig. 26 or 27 may be substituted for them or vice versa. The representation of the mutual impedance, Z_M, by an equivalent network (Fig. 25) makes it possible to represent the transformer of Fig. 23 by a T or π network containing only self impedances. This affords a great simplification in the analysis of filter circuits containing pairs of coils having mutual impedance between them in that it permits such circuits to be reduced to an equivalent series-shunt (or lattice or bridged-T) type structure. Consequently, the methods of design which have been built up for the series-shunt and kindred type structures may be directly applied to the solution of circuits containing such pairs of coils.

Two-Terminal Equivalent Meshes. A list of equivalent two-terminal reactance meshes, due to Zobel, has been given in Fig. 17. All of the meshes in Figs. 17B, C and D contain two inductance elements. Mutual inductance may exist between any two inductive elements without changing fundamentally the nature of the reactance meshes. This means that when mutual inductance exists between two coils in

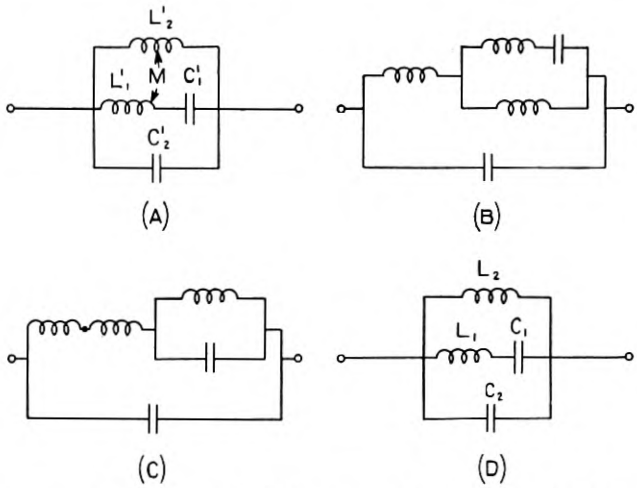


Fig. 28—Equivalent Two-Terminal Reactance Networks, Only One of Which Contains Mutual Inductance

any of these meshes, the mesh may be designed to be electrically equivalent to, and consequently can be substituted for, a corresponding mesh of the same type having no mutual inductance.

For example, consider the mesh shown in Fig. 28A which is potentially equivalent to the first reactance mesh of Fig. 17C and, consequently, to the other three reactance meshes of the same figure. The inductance elements L_1' and L_2', together with the mutual inductance M acting between them, may be represented by an equivalent T network, as previously stated. The reactance mesh formed by L_1', L_2', and M, together with its equivalent T and π forms, is shown in Fig. 29. By means of the relations given in Figs. 29A and B, it is possible to derive, from the structure of Fig. 28A, the equivalent structure shown in Fig. 28B. Likewise, from formulae (45) and (46) for the equivalence of the two structures of Fig. 17B, the mesh of Fig. 28C can be obtained from that of Fig. 28B. Furthermore, if the two inductances shown in series in Fig. 28C are merged, it is again possible, by means of the conversion formulae for the two meshes of

Fig. 17B, to determine the constants of the mesh shown in Fig. 28D from the known values of the constants of the structure of Fig. 28C.

The relations which must exist if the structure of Fig. 28D is to be equivalent to the structure shown in Fig. 28A, or vice versa, are given by the following relations

$$C_2 = C_2', \; L_1 = \frac{L_2'(L_1'L_2' - M^2)}{(L_2' \pm M)^2}, \tag{75}$$

$$L_2 = L_2', \;\; C_1 = C_1' \left(\frac{L_2' \pm M}{L_2'}\right)^2. \tag{76}$$

The upper and lower of the alternative signs, in the preceding equations, correspond respectively to series aiding and opposing connections. The equivalence of these four-element meshes makes it possible

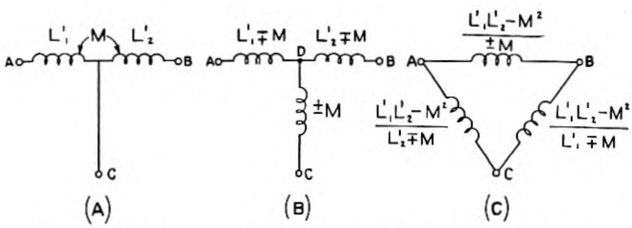


Fig. 29—Equivalent Three-Terminal Inductance Networks

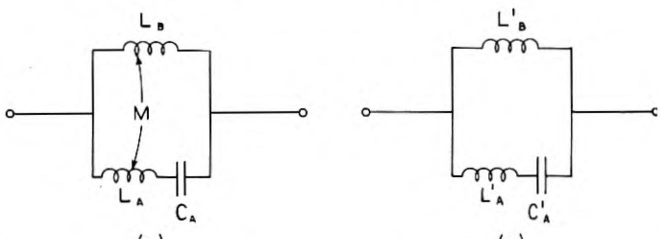


Fig. 30—Equivalent Two-Terminal Reactance Networks, Only One of Which Contains Mutual Inductance

to derive at once, the relations which must exist between certain equivalent three-element meshes involving mutual inductance. For example, if the capacity C_2' of Fig. 28A is zero, the mesh reduces to the three-element mesh of Fig. 30A and the formulae given above are then applicable for the equivalence of the structures of Figs. 30A and B.

In the same way that the meshes illustrated in Fig. 28 were shown to be potentially equivalent to each other, it is possible to prove that the meshes of Fig. 31 are potentially equivalent. The equivalence of the mesh shown in Fig. 31B to that of Fig. 31A is satisfied by the relations given in Figs. 29A and B. The equivalence of the mesh of Fig. 31C to that of Fig. 31B is governed by the equations (56 to 64) for the equivalence of the first and last structures of Fig. 17D. Fin-

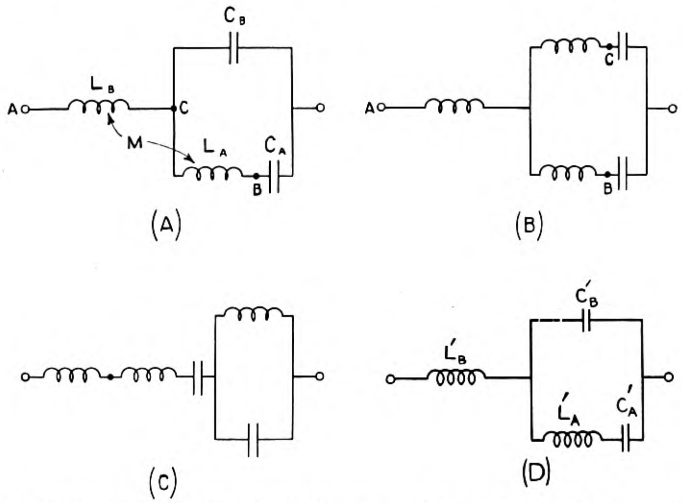


Fig. 31—Equivalent Two-Terminal Reactance Networks, Only One of Which Contains Mutual Inductance

ally, the equivalence of the mesh of Fig. 31D to that of Fig. 31C is controlled by the relations for the equivalence of the first two structures of Fig. 17D.

The formulae relating the constants of the structure shown in Fig. 31D to the corresponding constants of the structure shown in Fig. 31A are as follows:

$$L_A' = L_1\left(1+\frac{C_1}{C_2}\right)^2,\ L_B' = L_2,\ C_A' = \frac{C_2^2}{C_1+C_2},\ C_B = \frac{C_1C_2}{C_1+C_2}, \tag{77}$$

in which—

$$C_1 = \frac{C_AC_B(C_A+C_B)L_A^2}{[C_A(L_A\pm M)\pm MC_B]^2},\ C_2 = C_A+C_B, \tag{78}$$

and

$$L_1 = \frac{[C_A(L_A\pm M)\pm MC_B]^2}{(C_A+C_B)^2L_A},\ L_2 = \frac{L_AL_B-M^2}{L_A}. \tag{79}$$

The upper and lower of the alternative signs, in the preceding equations correspond, respectively, to series aiding and opposing connections.

The equivalence of these four-terminal meshes makes it possible to derive the relations which must exist for corresponding equivalent three-element meshes, with and without mutual inductance. For example, if in Fig. 31A, the capacity C_A is of infinite value, the mesh reduces to that shown in Fig. 32A and the formulae given above are applicable for the equivalence of the meshes of Figs. 32A and B.

The remaining meshes of Figs. 17C and D have similar potential equivalence to meshes of the same fundamental type but having mutual inductance between the respective pairs of coils.

Three-Terminal Equivalent Meshes. Three terminal meshes containing mutual inductance will now be discussed. It has been shown

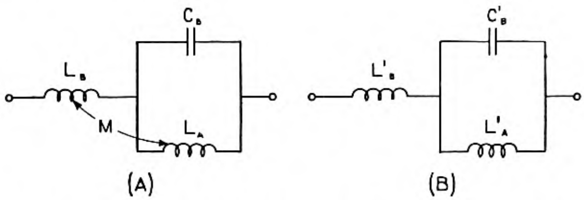


Fig. 32—Equivalent Two-Terminal Reactance Networks, Only One of Which Contains Mutual Inductance

that two coils, with mutual inductance between them (Fig. 29A), are equivalent to certain T and π structures containing only tangible inductances (Figs. 29B and C). Referring to Fig. 29B, it is seen that two coils, with series opposing mutual inductance between them (corresponding to the upper alternative signs in Fig. 29B), are equivalent to a T network having three positive inductance arms, provided the mutual inductance M is less than L_1' and L_2'. The values of these arms are respectively, $L_1'-M$, $L_2'-M$, and M. If M is larger than L_1', one arm of the equivalent T network is a negative inductance while the other two arms are positive inductances. Similarly, if M is larger than L_2', a different arm of the T network will be a negative inductance while the two remaining arms will be positive inductances. It is physically impossible for the value of M to be greater than both L_1' and L_2'. Hence, it is impossible for more than one arm of the T network, shown in Fig. 29B, to be a negative inductance.

When two coils have series aiding mutual inductance between them (the lower of the alternative signs in Fig. 29B) they are equivalent to a T network in which two of the arms consist of positive inductances viz., $L_1'+M$ and $L_2'+M$, while the third arm consists of a negative inductance of the value $-M$.

Whenever, in an equivalent T network, one of the arms is a positive (or negative) inductance, a corresponding arm of the π network will also be a positive (or negative) inductance. Consequently, as in the case of the equivalent T network, the equivalent π network shown in Fig. 29C may consist of three positive inductances or two positive inductances and one negative inductance, depending upon the sign and magnitude of M.

It is interesting to note that, in Fig. 29B, point D is in reality a concealed terminal, i.e., it cannot be regarded as physically accessible. There are, therefore, only three accessible terminals to the equivalent

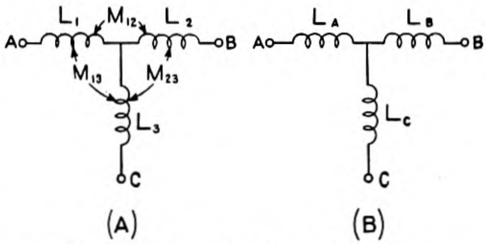


Fig. 33—Equivalent T Networks of Inductance

T network. In the π network shown in Fig. 29C there is no such concealed point. There are, however, as in the preceding case, three accessible terminals A, B and C.

When the mutual inductance, M, is equal to either one of the self inductances, L_1' (or L_2'), and the windings are connected in series opposing, the equivalent T and π networks of the transformer coalesce to the same L type network. For example, if $L_1'=M$ in Fig. 29A both the T and the π networks of Figs. 29B and C resolve into an L network whose vertical arm has the value M and whose horizontal arm is $L_2'-M$.

A problem of practical importance is the equivalence of T and π meshes, containing three coils with mutual inductance between all of the elements, to similar T and π meshes containing no mutual inductance. The T networks of Fig. 33 are potentially equivalent. The formulae governing their equivalence are

$$L_A=L_1+M_{12}+M_{13}-M_{23}, \tag{80}$$

$$L_B=L_2+M_{12}-M_{13}+M_{23}, \tag{81}$$

$$L_C=L_3-M_{12}+M_{13}+M_{23}. \tag{82}$$

In the above formulae, the signs correspond to the case of a series aiding mutual inductance between all the pairs of coils. When the mutual inductance between any two coils changes sign, the signs accompanying that mutual inductance in the above formulae are reversed.

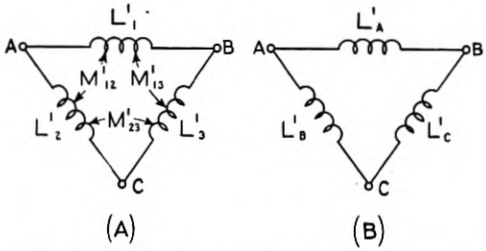


Fig. 34—Equivalent π Networks of Inductance

Similarly, the π networks of Fig. 34 are also potentially equivalent The formulae governing their equivalence are

$$L_A' = \frac{L_x L_y + L_x L_z + L_y L_z}{L_y}, \tag{83}$$

$$L_B' = \frac{L_x L_y + L_x L_z + L_y L_z}{L_z}, \tag{84}$$

$$L_C' = \frac{L_x L_y + L_x L_z + L_y L_z}{L_x}, \tag{85}$$

in which—

$$L_x = \frac{L_A'' L_B''}{L_A'' + L_B'' + L_C''} \mp M_{12}', \tag{86}$$

$$L_y = \frac{L_B'' L_C''}{L_A'' + L_B'' + L_C''} \mp M_{23}', \tag{87}$$

$$L_z = \frac{L_A'' L_C''}{L_A'' + L_B'' + L_C''} \mp M_{13}', \tag{88}$$

where

$$L_A'' = L_1' \pm M_{12}' \pm M_{13}', \tag{89}$$

$$L_B'' = L_2' \pm M_{12}' \pm M_{23}', \tag{90}$$

$$L_C'' = L_3' \pm M_{13}' \pm M_{23}'. \tag{91}$$

As in the preceding case, the upper of the two signs occurs with the series aiding mutual inductance between all the pairs of coils. When the mutual inductance between any two coils changes sign, the signs accompanying that mutual inductance in the above formulae are reversed.

At least two of the three inductances (in Fig. 33B or in Fig. 34B) will always be positive in sign while the third inductance may be either positive or negative. Consequently, three coils having mutual inductance between each of them and having only three accessible terminals offer no greater possibilities than do two coils having mutual inductance between them and having three terminals. In both cases the structure is equivalent to a T or π mesh composed of three self

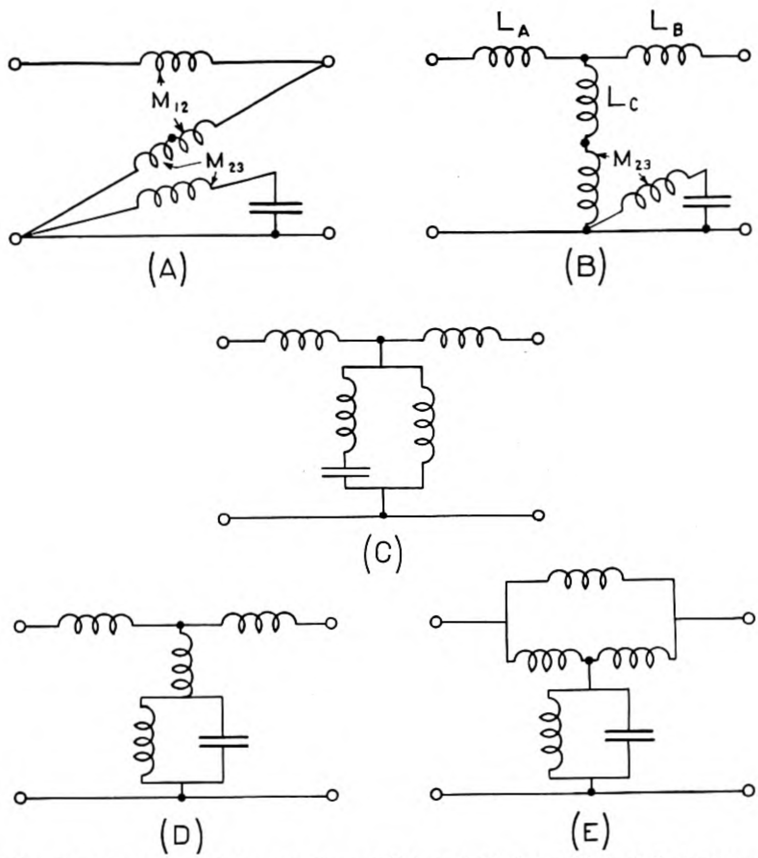


Fig. 35—Equivalent Filter Sections, With and Without Mutual Inductance

inductances, at least two of which must be positive. With specific relations between the various self and mutual inductances, it is possible for the three coils with mutual inductance between each of them to be equivalent (as in the case of two coils with mutual inductance) simply to an L network composed of two positive self inductances.

Since either two or three coils with mutual inductance between them are, in general, equivalent, at all frequencies, to a T or π network composed of three self inductances, it is possible to substitute the one type of mesh for the other in any kind of a circuit without affecting the currents or voltages external to the meshes involved. This substitution is always physically possible provided none of the arms of the equivalent T or π networks is a negative inductance.

The structures shown in Fig. 35 are illustrative of the power of equivalent networks as tools for the solution of filter structures containing mutual inductance. The equivalence of the structure shown in Fig. 35B to that of Fig. 35A is evident from the equivalence of two coils (Fig. 29) with mutual inductance (M_{12}) between them to three inductances, L_A, L_B and L_C without mutual inductance. Likewise,

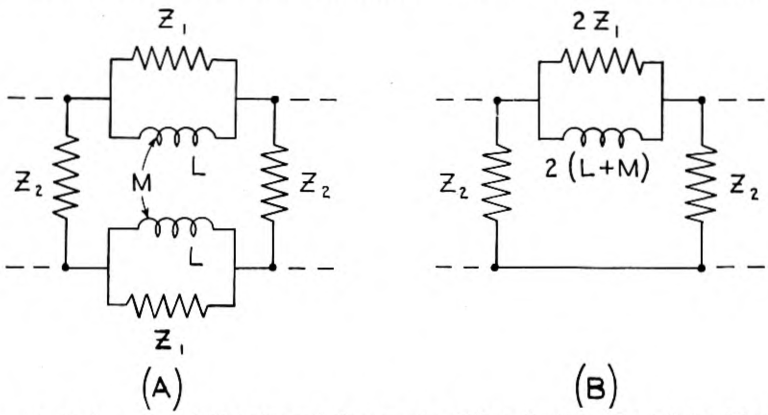


Fig. 36—Balanced and Unbalanced Forms of a Filter Section, Containing Mutual Inductance

the equivalence of the structure shown in Fig. 35C to that of Fig. 35B is obtainable by successive mesh substitutions. The equivalence of the structures shown in Fig. 35D and E to that of Fig. 35C are also obtainable from equivalences previously referred to. If the propagation and impedance characteristics of either of the structures of Fig. 35C or D are known, then the other structures shown in Fig. 35 will have the same characteristics. Furthermore, if the values of the constants of any one of these structures are known, the constants of any of the other structures are readily obtainable by means of transformation formulae.

In a large number of wave filters, the structures are unbalanced; that is, all of the series impedances are placed in one of the two line wires while the remaining wire is a short circuit. Ordinarily, the object in using such an unbalanced structure is to minimize the number of elements required in the series arms. It should be noted, however, (Fig. 36) that in case an inductance element enters into both series arms, it can be replaced, in symmetrical structures, by two equal windings of a single coil having mutual inductance between them and of such value that the series aiding inductance of these two coils is equal to the total inductance required in the corresponding unbalanced structure. For example, the structures shown in Figs. 36A and B are electrically equivalent to each other, that is, they have the same image impedance and transfer constant.

Types of Sections Obtainable Whose Equivalent Series-Shunt Sections Contain No Negative Inductances. It has previously been stated that an infinite number of types of series-shunt filter sections may be had, if no limitations are placed on the complexity of their reactance arms. It has also been stated, however, that for filters employing only one transmission or one attenuation band, the maximum number of elements which can ordinarily be used economically per section is six. A similar limitation exists when mutual inductance is employed, in that sections can seldom be economically used whose prototype structures contain more than six reactance elements.

Inasmuch as by the equivalences which have been discussed, many varient forms of a section may exist, which forms are reducible to the same series-shunt prototype, an effort only to list and discuss the prototype sections will be made. The prototype to which any given section then reduces will readily be found by the application of the foregoing principles. A few examples will later serve to make this clear.

In considering the prototype sections which exist when mutual inductance is present in a filter section, we shall first list the reactance meshes of which mutual inductance may form a part. Referring to Fig. 5, an inspection of the equivalences so far discussed will show that the following meshes may be partly or wholly composed of mutual inductance:

1, 3, 4, 5 (*a* and *b*), 7 (*a* and *b*), and 8 (*a* and *b*).

Consequently, a large number of the sections listed in Table II and formed from the reactance meshes of Fig. 5 may represent *not only actual sections containing no mutual inductance, but also equivalent prototypes of sections containing mutual inductance.* Sections containing mutual inductance within only the series arm or the shunt arm, respectively, are not included in this discussion since such arms may be readily reduced to equivalent arms, without mutual inductances, by the substitution of equivalent two-terminal meshes. The prototypes which are under discussion are listed below:

Low pass	*High pass*
$1-3$, $5-3$	$4-1$, $4-5$

Band pass

3−1, 1−4, 3−3, 4−4, 1−5, 5−1, 3−7, 3−5, 8−4, 4−8, 5−4, 5−5, and 7−3.

Sections corresponding to the equivalent series-shunt prototypes listed will have the same impedance and propagation characteristics as the prototype, and may be used indiscriminately in place of the prototype. Consequently, when a section has been reduced to any of the above prototypes, its various characteristics may be found from Table II and Figs. 7 and 8.

As an example of structures which have mutual inductance and which are equivalent to structures listed above, consider the section

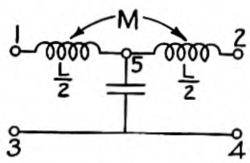


Fig. 37—Low Pass Filter Section Containing Two Coils, Having Mutual Inductance Acting Between Them, and a Condenser Shunted From Their Junction Point

shown in Fig. 37. This section contains two coils having mutual inductance, and a condenser shunted from their junction point. The three-terminal mesh formed by the two coils $L/2$ and $L/2$, together with their series opposing mutual inductance M, may be represented, as in Fig. 29B, by its equivalent T mesh. The resulting equivalent section is that shown in Fig. 38. The structure of Fig. 38, having a series reactance mesh corresponding to No. 1 of Fig. 5, and a shunt

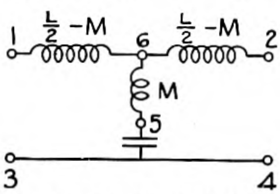


Fig. 38—Filter Section Containing No Mutual Inductance, Equivalent to the Section of Fig. 37

reactance mesh corresponding to No. 3 of Fig. 5 is that listed as 1−3 in Table II and in the above list. Consequently, it has propagation characteristic No. 2 of Fig. 7, and mid-series image impedance characteristic No. 1 of Fig. 8. The section of Fig. 37 may, consequently, be joined at either end to any structure having a mid-series image impedance characteristic such as that designated as characteristic No. 1 of Fig. 8. The section of Fig. 37 is not capable of mid-shunt termination since point 6 of Fig. 38 is not physically accessible.

Similarly, the section shown in Fig. 39 is equivalent to the series-shunt structure of Fig. 40. If the transformer mesh in Fig. 39, formed by $2L_2$, M and $2L_2$ be replaced by its equivalent π mesh,—assuming series opposing windings—the structure of Fig. 40 results.

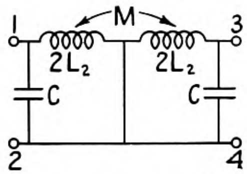


Fig. 39—Band Pass Filter Section Containing Mutual Inductance

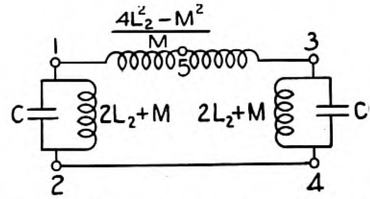


Fig. 40—Filter Section, Containing No Mutual Inductance, Equivalent to the Section of Fig. 39

This structure is listed as band pass section $1-4$ in Table II and has propagation characteristic No. 7 of Fig. 7, and mid-shunt image impedance characteristic No. 14 of Fig. 8. Consequently, the section of Fig. 39 may be joined efficiently to any filter section of Table II having the mid-shunt image impedance characteristic No. 14 of Fig. 8 or to any section containing mutual inductance and having the same mid-shunt image impedance characteristic. The section of Fig. 39 is not capable of mid-series termination, since point 5 of inductive element $1-3$ of Fig. 40 is not physically accessible.

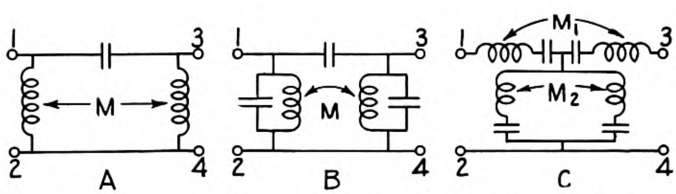


Fig. 41—Examples of Filter Sections Containing Mutual Inductance

Three further examples of the substitutions which have been discussed are represented in Figs. 41A, B, and C. By means of substitutions these structures are evidently equivalent to series-shunt sections $4-1$ (mid-shunt terminated), $4-4$, (mid-shunt terminated), and $3-7$ (mid-series terminated), respectively, and they have the characteristics detailed in Table II. The above examples represent only a few of the many variant forms of structures which may be constructed by means of the various equivalences heretofore discussed.

The representation of the characteristics of the structures of Table III is similar to the scheme of Table II. The figures at the top and side (for example $1-3'$) indicate respectively, the series and shunt reactance meshes of Figs. 5 and 42 which form the prototype sections.

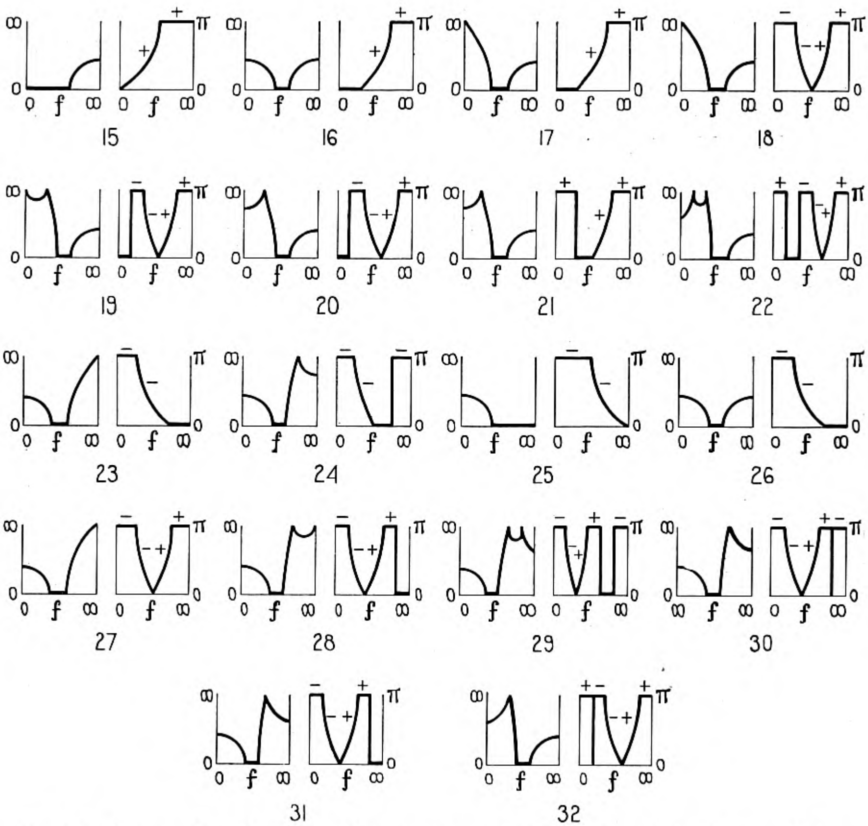


Fig. 44—Propagation Constant (Attenuation and Phase Constant) Characteristics of Filter Sections Containing Negative Inductances, Shown in Symbolic Form

The figures in the corresponding box (for example, $15-1-*$) indicate that the structure has propagation characteristic No. 15 of Fig. 44, and mid-series image impedance No. 1 of Fig. 8. The symbol * indicates, when inserted in the second or third position, that the structure is not physically capable of mid-series or mid-shunt termination, respectively.

It will be noted that only one low pass prototype section $(1-3')$ is given in the table, exclusive of special cases of band filter structures.

Its attenuation characteristic (No. 15 of Fig. 44) is unique as a low pass characteristic in that *the attenuation constant is finite at all frequencies.* The phase characteristic simulates, in a general way, that of the two element low pass filter (see propagation characteristic No. 1 of Fig. 7) but the phase shift in the transmission band is, in general, different. Since the structure has mid-series image impedance characteristic No. 1 it may be joined efficiently (i.e., without reflection losses) to sections of the $1-2$ and $1-3$ types.

Similarly, high pass prototype section $4'-1$ has a unique high pass attenuation characteristic in that the attenuation constant is finite at all frequencies. The phase characteristic is, in general, similar to that of the two element high pass filter $2-1$ except for the values of the phase constant in the transmission band. The section may be joined efficiently at mid-shunt to sections of the $2-1$ and $4-1$ types—since it has the same mid-shunt image characteristic (No. 9).

The attenuation characteristics of the band pass prototypes listed in Table III will, in general, differ from the attenuation characteristics of structure listed in Table II. However, many of them differ only in minor respects and could have been represented identically in the symbolic fashion of Fig. 7. Inasmuch as such structures will not, however, have exactly the same attenuation characteristics for given cut-off frequencies and frequencies of infinite attenuation, different symbols or diagrams have been employed to represent them.

Certain characteristics are worthy of comment because they are not obtainable, even approximately, in structures not having negative inductance. For example, propagation characteristics Nos. 16 and 26 (Fig. 44) are band pass filter characteristics having finite attenuation at all frequencies. Characteristics No. 22 and No. 29 are unique in that there exist two frequencies of infinite attenuation, located on one side of the pass band. The attenuation constant is, in general, finite at zero and at infinite frequencies. Characteristics 19 and 28 are special cases of Nos. 22 and 29, respectively, and have two frequencies of infinite attenuation on one side of the pass band. In the case of 19, the attenuation is infinite at zero frequency and at a frequency between zero and the lower cut-off frequency. Characteristic 28 has infinite attenuation at infinite frequency and also at a frequency between the upper cut-off frequency and infinite frequency. Characteristics Nos. 18 and 27 have confluent band characteristics and have only one frequency of infinite attenuation, located either at zero frequency or at infinite frequency. Finally, characteristics Nos. 20 and 31 are confluent characteristics in each of which one frequency of infinite attenuation occurs and the attenuation is finite at zero frequency and infinite frequency.

As a general rule the phase shift characteristics shown in Fig. 44 are similar to the corresponding characteristics shown in Fig. 7. The phase characteristics of the former, within the pass bands are, in general, however, of a distinctly different character than those of the latter even though the phase constant at the cut-off frequency and the mid-frequency may be the same. Phase characteristics 21 and 24 (Fig. 44) are of special interest, however, in that while they belong to the peak type sections, the phase is of the same sign throughout the entire frequency range. Also phase characteristics 22, 29, 30 and 32 have a unique property, for band pass structures, in that the phase undergoes a change in sign within one attenuation band.

In regard to the impedance characteristics, it is noted from Table III that *no novel impedance characteristics are obtained in structures having negative inductances as compared to the structures not having negative inductances.* This is a valuable property of the prototype structures listed in Table III as it permits composite filters to be readily formed utilizing both the sections of Tables II and III.[16]

Characteristics of a Typical Filter. In order to illustrate the derivation of design formulae for a specific prototype having negative inductances, consider as an example the band pass structure $3-3'$ of Table III. We shall neglect the effect of dissipation on the characteristics of the structure, as the treatment of dissipation has been previously outlined. The prototype cited is illustrated in Fig. 45A. Two

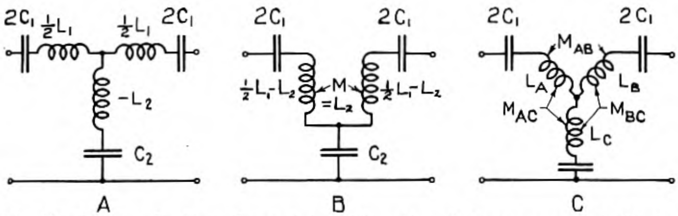


Fig. 45—Prototype Section Containing Negative Inductance, and Two of Its Physically Realizable Forms

methods of physically obtaining such a prototype are illustrated in Figs. 45B and C. In this structure the series impedance Z_1 is

$$Z_1 = j\left(\omega L_1 - \frac{1}{\omega C_1}\right). \tag{92}$$

[16] For a general method of proving the equality of the image impedances of sections containing negative inductance and of appropriate sections containing no negative inductance, refer to the Appendix.

The impedance of the shunt arm is

$$Z_2 = -j\left(\omega L_2 + \frac{1}{\omega C_2}\right). \tag{93}$$

The ratio, $Z_1/4Z_2$, which controls the attenuation and phase constants, per section, of the structure is

$$\frac{Z_1}{4Z_2} = \frac{j\left(\omega L_1 - \frac{1}{\omega C_1}\right)}{-j\left(\omega L_2 + \frac{4}{\omega C_2}\right)} = \frac{C_2}{4C_1}\,\frac{1-L_1C_1\omega^2}{1+L_2C_2\omega^2}. \tag{94}$$

From the impedance characteristics of reactance meshes 3 and 3′, as illustrated in Figs. 6 and 43, and the combined reactance character-

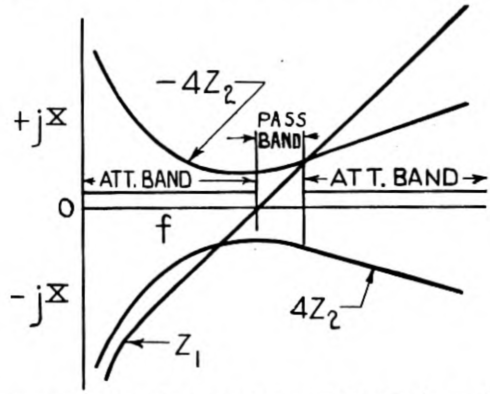


Fig. 46—Reactance-Frequency Characteristics of the Series and Shunt Arms of the Prototype Section of Fig. 45-A

istics of Fig. 46 for Z_1, $4Z_2$ and $-4Z_2$, it will be noted that the lower cut-off frequency, f_1, is that at which $Z_1=0$. Hence,

$$f_1 = \frac{1}{2\pi\sqrt{L_1C_1}}. \tag{95}$$

Similarly, the upper cut-off frequency is that at which $Z_1=-4Z_2$ or $j\omega L_1 - j/\omega C_1 = j4\omega L_2 + j4/\omega C_2$. From this relationship, the upper cut-off frequency is

$$f_2 = \frac{1}{2\pi}\sqrt{\frac{C_2+4C_1}{C_1C_2(L_1-4L_2)}}. \tag{96}$$

Let f_r be assumed as the frequency where Z_2 is a minimum, that is, where $\omega^2L_2C_2=1$. We may then write

$$f_r = \frac{1}{2\pi\sqrt{L_2C_2}}. \tag{97}$$

Substituting the above values of f_1, f_2 and f_r in formula (94) we obtain for $Z_1/4Z_2$

$$\frac{Z_1}{4Z_2}=\frac{1-\left(\frac{f}{f_1}\right)^2}{1+\left(\frac{f}{f_r}\right)^2}\,\frac{\left(\frac{f_2}{f_r}\right)^2+1}{\left(\frac{f_2}{f_1}\right)^2-1}. \tag{98}$$

From this last expression the attenuation and phase characteristics may be plotted from formulae (22) and (23) or from Figs. 11 and 12. The attenuation and phase constant characteristics are shown symbolically as characteristic 16 of Fig. 44. This structure has unusual attenuation properties which have already been discussed.

From equation (6) and the values of Z_1 and Z_2, in (92) and (93), the mid-series image impedance (Z_o), at the mid-frequency, is

$$Z_o=\frac{1}{2}\left[\sqrt{\frac{L_1}{C_1}+\frac{4L_1}{C_2}}-\sqrt{\frac{L_1}{C_2}-\frac{4L_2}{C_1}}\right]. \tag{99}$$

Since the mid-series image impedance, at any frequency, is the same as that of filter section $3-3$, we have:

$$Z_I=Z_o\sqrt{1-\frac{\left[\frac{f}{f_m}-\frac{f_m}{f}\right]^2}{\left[\frac{f_2}{f_m}-\frac{f_m}{f_2}\right]^2}}=Z_o\sqrt{1-\frac{\left[\frac{1}{\sqrt{f_1f_2}}-\frac{\sqrt{f_1f_2}}{f}\right]^2}{\left[\sqrt{\frac{f_2}{f_1}}-\sqrt{\frac{f_1}{f_2}}\right]^2}} \tag{100}$$

where f_m is the mid-frequency ($f_m=\sqrt{f_1f_2}$), as before.

The prototype is not capable of mid-shunt termination, hence, its hypothetical mid-shunt impedance characteristic will not be derived.

From the preceding formulae, explicit expressions may be derived for the values of L_1, C_1, L_2 and C_2

$$L_1=\frac{Z_om'}{\pi(f_2-f_1)}, \tag{101}$$

$$C_1=\frac{f_2-f_1}{4\pi f_1^2Z_om'}, \tag{102}$$

$$L_2=\frac{-Z_o}{\pi(f_2-f_1)}\,\frac{1-m'^2}{4m'}, \tag{103}$$

$$C_2=\frac{(f_2-f_1)m'}{\pi Z_o(f_2^2-f_1^2m'^2)}, \tag{104}$$

$$m'=\sqrt{1+\frac{\left(\frac{f_2}{f_1}\right)^2-1}{\left(\frac{f_r}{f_1}\right)^2+1}}. \tag{105}$$

As a numerical example of the solution of the prototype discussed assume, as in the example following equation (41), that the lower cut-off frequency f_1 is 20,000 cycles and that the upper cut-off frequency f_2 is 25,000 cycles. Assume f_r, a convenient parameter for the families of attenuation and phase constant curves which this section may have, for any given cut-off frequency, to be 30,000 cycles. Assume that the value of the mid-series image impedance Z_o at the mid-frequency is 600 ohms; then from formula (99) $m' = 1.083$: hence $L_1 = .0412$ henries, $C_1 = .00153 \times 10^{-6}$ farads, $L_2 = .00152$ henries and $C_2 = .0184 \times 10^{-6}$ farads. The structure with the numerical values of inductance and capacity for this specific example is shown in Fig. 47A.

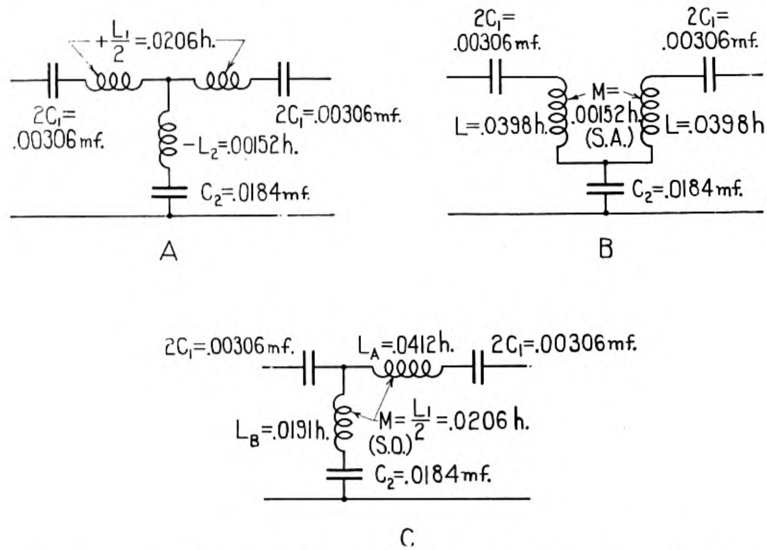


Fig. 47—Numerical Example of Equivalent Filter Sections Containing Negative Inductance

If, for the T mesh inductances in Fig. 47A, we substitute a transformer mesh having the values shown in Fig. 47 B, the mesh of the latter figure is electrically equivalent to the prototype structure and is an example of the method of employing the structure. Similarly, Fig. 47C illustrates the substitution of another type of three element mesh for the coil mesh of the prototype structure of Fig. 47A and is another example of the manner in which the prototype may be physically expressed.

The structure of Fig. 47B represents a similar case to that of 48A. However, as the mutual inductance is here series opposing, the prototype series-shunt equivalent structure is shown in Fig. 48B and contains no negative inductances. It will be found that the values chosen correspond to the numerical example of the structure 3−3 following equation 41.

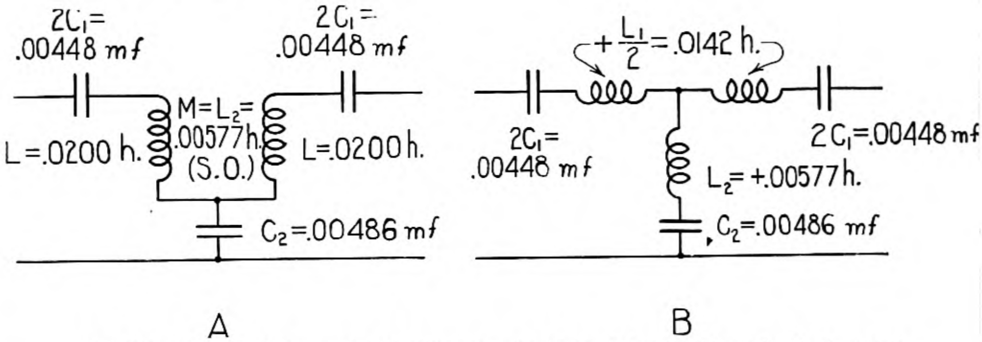


Fig. 48—Numerical Example of a Filter Section Containing No Negative Inductance

APPENDIX

Conditions for the Equality of the Image Impedances of Typical Filter Structures

It has been stated that the formation of recurrent and composite wave filters is dependent upon the maintenance of equal image impedance characteristics (of the sections or half-sections joined) at each junction point throughout the filter.

A general method of ascertaining the conditions for the equality of image impedance characteristics will be demonstrated by illustrations from typical pairs of sections.

Illustration No. 1—Negative Inductance in Shunt Arm of One Structure. Consider the filter sections listed as 3–4 (confluent structure) in Table II, and 3–1′ in Table III. It will be shown that, under proper conditions, their mid-series image impedance characteristics may be made equal at all frequencies. (By reference to the above tables, both sections have mid-series impedance characteristic No. 13 of Fig. 8).

From equation (6)

$$Z_I^2 = Z_1 Z_2 + \frac{Z_1^2}{4}. \tag{106}$$

In Fig. 49, let

$$Z_1 \equiv Z_{1A} + Z_{1B} = j\omega L_1 + \frac{1}{j\omega C_1}, \tag{107}$$

$$Z_1' \equiv K_A Z_{1A} + K_B \; Z_{1B}, \tag{108}$$

and $$Z_2' \equiv -K_C Z_{1A}, \tag{109}$$

where $K_A \equiv L_1'/L_1$, $K_B \equiv C_1/C_1'$ and $K_C \equiv L_2'/L_1$. (110)

From (106)

$$Z_I{}^2 = R^2 + \frac{Z_1{}^2}{4} \tag{111}$$

in which

$$R \equiv \sqrt{\frac{L_2}{C_1}} \equiv \sqrt{\frac{L_1}{C_2}}. \tag{112}$$

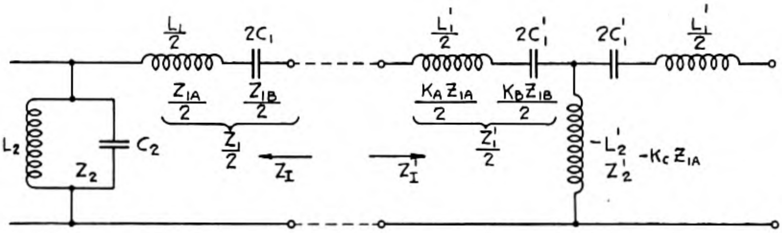


Fig. 49—Two Structures Having Equal Mid-Series Image Impedances, One of Which Contains a Negative Inductance in Its Shunt Arm

From (107) and (111)

$$Z_I{}^2 = R^2 + \tfrac{1}{4}(Z_{1A} + Z_{1B})^2 = 1/4 Z_{1A}^2 + (1 + K/2)R^2 + 1/4 Z_{1B}^2 \tag{113}$$

where $K \equiv Z_{1A} Z_{1B}/R^2 = L_1/L_2 = C_2/C_1$. (114)

Now from (106) and (108)

$$(Z_I')^2 = Z_1' Z_2' + \frac{(Z'_1)^2}{4} = \left(\frac{K_A{}^2}{4} - K_A K_C\right) Z_{1A}^2 + \left(\frac{K_A K_B}{2} - K_B K_C\right) K R^2 + \frac{K_B{}^2}{4} Z_{1B}^2. \tag{115}$$

Since, by postulation, in Fig. 49, $Z_1 = Z_1'$, we may equate the coefficients of (113) and (115). This gives

$$\frac{1}{4} = \frac{K_A{}^2}{4} - K_A K_C, \tag{116}$$

$$1 + \frac{K}{2} = \left(\frac{K_A K_B}{2} - K_B K_C\right) K, \tag{117}$$

and $$\frac{1}{4} = \frac{K_B{}^2}{4}. \tag{118}$$

Whence $$K_B \equiv \frac{C_1}{C_1'} = 1, \tag{119}$$

and $$K_A \equiv \frac{L_1'}{L_1} = \frac{L_1'C_1'}{L_1C_1} = \frac{f_M^2}{f_1^2} = \frac{f_2}{f_1}, \tag{120}$$

where f_1 and f_2 are the lower and upper cut-off frequencies, respectively, and $f_M \equiv \sqrt{f_1 f_2}$ of the structures of Fig. 49.

From (116) and (120)

$$K_C \equiv \frac{L_2'}{L_1} = \frac{1}{4}\left(K_A - \frac{1}{K_A}\right) = \frac{1}{4}\left(\frac{f_2}{f_1} - \frac{f_1}{f_2}\right). \tag{121}$$

Therefore, when the relationships between the constants of the two structures of Fig. 49 satisfy equations (119), (120) and (121), the structures will have the same mid-series image impedance characteristics. Explicit relations for the values of C_1', L_1' and L_2' may be obtained from equations (119), (120) and (121) as follows:

$$C_1' = C_1, \tag{122}$$

$$L_1' = L_1\frac{f_2}{f_1}, \tag{123}$$

$$L_2' = \frac{L_1}{4}\left(\frac{f_2}{f_1} - \frac{f_1}{f_2}\right). \tag{124}$$

Consequently, if the constants and cut-off frequencies of a confluent structure are known, the constants of a structure of the 3–1′ form having an identical mid-series image impedance characteristic can be derived from equations (122), (123) and (124).

Illustration No. 2—Negative Inductance in Series Arm of One Structure. Consider next the filter sections listed as 3–4 (confluent structure) in Table II and 1′–4 in Table III. It will be shown that, under proper conditions, their mid-shunt image impedance characteristics may be made equal at all frequencies. (By reference to the above tables, both sections have mid-shunt impedance characteristic No. 14 of Fig. 8).

From equation (7)

$$Y_I^2 = Y_1 Y_2 + \frac{Y_2^2}{4}, \tag{125}$$

where $Y_1 = 1/Z_1$, $Y_2 = 1/Z_2$ and $Y_I = 1/Z_I$.

In Fig. 50, let

$$Y_2 \equiv Y_{2A} + Y_{2B} = \frac{1}{j\omega L_2} + j\omega C_2. \tag{126}$$

$$Y_2' \equiv K_A Y_{2A} + K_B Y_{2B}, \tag{127}$$

and

$$Y_1' \equiv -K_C Y_{2A}, \tag{128}$$

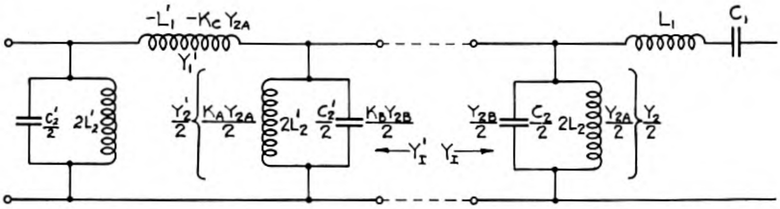


Fig. 50—Two Structures Having Equal Mid-Shunt Image Impedances, One of Which Contains a Negative Inductance in Its Series Arm

where

$$K_A \equiv L_2/L_2', \quad K_B \equiv C_2'/C_2 \text{ and } K_C \equiv L_2/L_1'. \tag{129}$$

From (125)

$$Y_I^2 = G^2 + \frac{Y_2^2}{4} \tag{130}$$

in which

$$G \equiv \sqrt{\frac{C_1}{L_2}} \equiv \sqrt{\frac{C_2}{L_1}} \tag{131}$$

From (126) and (130)

$$Y_I^2 = G^2 + 1/4(Y_{2A} + Y_{2B})^2 = 1/4\,Y_{2A}^2 + (1 + K/2)G^2 + 1/4\,Y_{2B}^2 \tag{132}$$

where

$$K \equiv Y_{2A} Y_{2B}/G^2 = L_1/L_2 = C_2/C_1. \tag{133}$$

Now from (125) and (127)

$$(Y_I')^2 = Y_1' Y_2' + \frac{(Y_2')^2}{4} = \left(\frac{K_A^2}{4} - K_A K_C\right) Y_{2A}^2 +$$

$$\left(\frac{K_A K_B}{2} - K_B K_C\right) K G^2 + \frac{K_B^2}{4} Y_{2B}^2 \tag{134}$$

Since, by postulation, in Fig. 50, $Y_I = Y_I'$, we may equate the coefficients of (132) and (134). This gives

$$\frac{1}{4} = \frac{K_A^2}{4} - K_A K_C, \tag{135}$$

$$1 + \frac{K}{2} = \left(\frac{K_A K_B}{2} - K_B K_C\right) K, \tag{136}$$

and $$\frac{1}{4}=\frac{K_B^2}{4}. \tag{137}$$

Whence $$K_B \equiv \frac{C_2'}{C_2}=1, \tag{138}$$

and $$K_A \equiv \frac{L_2}{L_2'}=\frac{L_2C_2}{L_2'C_2'}=\frac{f_2^2}{f_M^2}=\frac{f_2}{f_1} \tag{139}$$

where f_1 and f_2 are the lower and upper cut-off frequencies, respectively, and f_M is the mean frequency $(\sqrt{f_1 f_2})$ of the structures of Fig. 50.

From (135) and (139)

$$K_C \equiv \frac{L_2}{L_1'}=\frac{1}{4}\left(K_A-\frac{1}{K_A}\right)=\frac{1}{4}\left(\frac{f_2}{f_1}-\frac{f_1}{f_2}\right). \tag{140}$$

Therefore, when the relationships between the constants of the two structures of Fig. 50 satisfy equations (138), (139) and (140), the structures will have the same mid-shunt image impedance characteristics. Explicit relations for the values of C_2', L_2' and L_1' may be obtained from equations (138), (139) and (140) as follows:

$$C_2'=C_2, \tag{141}$$

$$L_2'=L_2\frac{f_1}{f_2}, \tag{142}$$

$$L_1'=\frac{4L_2}{\left(\frac{f_2}{f_1}-\frac{f_1}{f_2}\right)}. \tag{143}$$

Therefore, if the constants and cut-off frequencies of a confluent structure are known, the constants of a structure of the $1'-4$ form having an identical mid-shunt image impedance characteristic can be derived from equations (141), (142) and (143).

BIBLIOGRAPHY

1. Thévenin, M. L., "Sur un Nouveau Théorème d'Electricité Dynamique," *Comptes Rendus*, Vol. 97, pp. 159–161 (1883).
2. Kennelly, A. E., "The Equivalence of Triangles and Three-Pointed Stars in Conducting Networks," *Electrical World and Engineer*, New York, Vol. XXXIV, pp. 413–414, Sept. 16, 1899.
3. Campbell, G. A., "Cisoidal Oscillations," *Trans. A. I. E. E.*, Vol. XXX, Part II, pp. 873–909 (1911).
4. Campbell, G. A., U. S. Patents Nos. 1,227,113 and 1,227,114 (1917).
5. Gherardi, B. and Jewett, F. B., "Telephone Repeaters," *Trans. A. I. E. E.*, 1919.
6. Wagner, K. W., *Arch. fur Elektrotechnik*, Vol. 8, p. 61 (1919); *E. T. Z.*, Aug. 7, 1919.

7. Van der Bijl, H. T., "Thermionic Vacuum Tubes," Published 1920.
8. Pierce, G. W., "Electric Oscillations and Electric Waves," Published, 1920.
9. Colpitts, E. H. and Blackwell, O. B., "Carrier Current Telephony and Telegraphy," *Trans. A. I. E. E.*, Feb., 1921.
10. Clement, L. M., Ryan, F. M., and Martin, D. K., "The Avalon-Los Angeles Radio Toll Circuit," *Proc. I. R. E.*, May, 1921.
11. Fletcher, H., "The Nature of Speech and Its Interpretation," *Jour. Franklin Inst.*, June, 1922.
12. Campbell, G. A., "Physical Theory of the Electric Wave-Filter," *Bell Sys. Tech. Jour.*, Nov., 1922.
13. Zobel, O. J., "Theory and Design of Uniform and Composite Electric Wave-Filters," *Bell Sys. Tech. Jour.*, Jan., 1923.
14. Rose, A. F., "Practical Application of Carrier Telephone and Telegraph in the Bell System," *Bell Sys. Tech. Jour.*, April, 1923.
15. Hartley, R. V. L., "Relation of Carrier and Side-Bands in Radio Transmission," *Bell Sys. Tech. Jour.*, April, 1923.
16. Bown, C. D., Englund, C. R., and Friis, H. T., "Radio Transmission Measurements," *Proc. I. R. E.*, April, 1923.
17. Peters, L. J., "Theory of Electric Wave Filters Built up of Coupled Circuit Elements," *Jour. A. I. E. E.*, May, 1923.
18. Demarest, C. S., "Telephone Equipment for Long Cable Circuits," *Bell Sys. Tech. Jour.*, July, 1923.
19. Nichols, H. W. and Espenschied, L., "Radio Extension of the Telephone System to Ships at Sea," *Bell Sys. Tech. Jour.*, July, 1923.
20. Carson, J. R. and Zobel, O. J., "Transient Oscillations in Electric Wave-Filters," *Bell Sys. Tech. Jour.*, July, 1923.
21. Arnold, H. D. and Espenschied, L., "Transatlantic Radio Telephony," *Bell Sys. Tech. Jour.*, Oct., 1923.
22. Best, F. H., "Measuring Methods for Maintaining the Transmission Efficiency of Telephone Circuits," *Jour. A. I. E. E.*, Feb., 1924.
23. Casper, W. L., "Telephone Transformers," *Jour. A. I. E. E.*, March, 1924.
24. Slaughter, N. H. and Wolfe, W. V., "Carrier Telephony on Power Lines," *Jour. A. I. E .E.*, April, 1924.
25. Foster, R. M., "A Reactance Theorem," *Bell Sys. Tech. Jour.*, April, 1924.
26. Martin, W. H., "The Transmission Unit and Telephone Transmission Reference System," *Bell Sys. Tech. Jour.*, July, 1924.
27. Zobel, O. J., "Transmission Characteristics of Electric Wave-Filters," *Bell Sys. Tech. Jour.*, Oct., 1924.

Some Contemporary Advances in Physics—VI
Electricity in Gases

By KARL K. DARROW

1. Introduction

THE physicists of a quarter of a century ago, who devoted themselves to the study of electricity in gases, were happily inspired; for among the myriad of intricate and obscure phenomena which they observed there are some few of an extreme simplicity, in which the qualities of the individual atoms of matter and electricity are manifest; in analyzing these they entered upon the path that led most directly to the deeper understanding of nature which is superseding the physics of the nineteenth century, and the physics of today is founded upon their efforts. The electron was perceived for the first time in the course of observations on the electric discharge in rarefied gases, and other experiments in the same field established the atom in science as a real and definite object. The discovery of the atom is commonly credited to the chemists; yet fifteen years have not passed since students of chemistry were being warned by a famous teacher that "atom" and "molecule" are figurative words, not on any account to be taken literally! The laws of chemical combination were held insufficient to prove that atoms have any real existence; though elements may always combine with one another in unchanging proportions, this does not prove anything about the weights of the atoms, or their sizes, or their qualities, or even that all the atoms of an element have the same weight, or even that there are any atoms at all. Now that we are past the necessity for this caution, and can count atoms, and measure their masses, and infer something about their structure, and estimate how close together they can approach, and know what happens to them when they strike one another or are struck by electrons; now that we can fill in the picture of the atom with so many and so diverse details, we are indebted for this progress chiefly to the men who gathered the data and made the theories concerning the conduction of electricity in gases. Many will remember how in the years before the great war this field of research seemed the most vital part of physics, the most inspired with a sense of new life and swift advance; now others share with it the centre of the stage, but they won their places chiefly because of the light it shed upon them.

It seems strange that the flow of electricity in gases should have proved easier to interpret than the flow of electricity in metals, which in appearance is certainly by far the simpler. One applies the terminals of a battery to the ends of a wire, and promptly the electric potential distributes itself with a uniform gradient along the wire and a current flows steadily down it. So rigorously is the current proportional to the voltage between the ends of the wire, over very wide ranges of voltage and current, that we regard the ratio as an essential constant of the wire; and we regard the ratio of potential-gradient (electric field) to current density as an essential characteristic of the metal, and give it a name—resistivity or specific resistance—and refer to theories of conduction in metals as theories of metallic resistance. It all seems exceedingly simple, and yet in the foregoing article of this series I have shown how all the attempts to interpret it have gone in vain. Much more complex in appearance is the discharge through a gas. One applies the terminals of a battery to a pair of electrodes facing one another in the open air, and perhaps nothing happens, or so minute a current flows that the most delicate of instruments is demanded to detect it; and then when the battery-voltage is very slightly raised, there may be an explosion with a blaze of light, dissociating the gas and corroding the electrodes, and draining off the available electricity in a moment. Or if one of the electrodes is acutely pointed there may be glows and luminous sheaths around it or tentacles of bluish light ramifying from it far and wide through the air. Or the discharge may rise to the heat of incandescence, and the gas and the electrodes shine with a blinding radiance, the brightest light that can be kindled on the earth. Or if the electrodes are enclosed in a tube containing a rarefied gas or vapor, the gas flares up into an extraordinary pattern of light and shade, lucent vividly-colored clouds floating between regions glowing feebly or obscure; and as the gas is gradually pumped away, the pattern changes and fades, a straight beam of electrons manifests itself by a luminous column traversing the tube, the glass walls flash out in a green fluorescence, and finally all becomes extinct. As for that even gradient, and that constant proportion between current and field strength distinguishing the metals, we cannot find them here. There is no such thing as the resistance of a gas; we had better forget the word, we cannot attach any physical meaning to the ratio of current and voltage.

I must not give the impression that all these manifold forms of the electric discharge in gases are understood. Certain of the simplest of them have been clarified, and as a result still simpler ones have been realized and comprehended in their turns, and so on down to the simplest of all, which is the discharge across a vacuum. This sounds somewhat like a paradox and so it would have seemed thirty or forty years ago, when electricity was thought to be inseparable from matter, and the only known discharges across gases were the discharges in which the gas plays an indispensable role. It is important to note the manner of this evolution, for much of the history of modern physics is dominated by it. We should not be nearly so far advanced as we are, had we not learned two things: how to reduce the amount of gas in a tube until an electron can fly clear across it with scarcely any chance of meeting an atom, and how to persuade an electron to emerge from a metal otherwise than by starting a discharge in a gas over its surface. We who are so familiar with the idea of electrons boiling out of a hot wire, or driven out of a cold metal plate by light shining upon it, or fired as projectiles out of exploding atoms, find it difficult to imagine the confusion which of necessity prevailed when all these processes were unknown. In the early stages of research into the discharge in gases, it was made clear that of each self-maintaining discharge a stream of electrons flowing out of the negative electrode is an essential part; the electron-stream maintains the gas-discharge, and reciprocally the gas-discharge maintains the electron-stream. The latest stage commenced when it was made possible to produce and maintain such an electron-stream independently of any gas-discharge, and deal with it at will.

Let me then begin the exposition with this idea, which so many years of research were required to render acceptable: the idea of a stream of electrons emerging from a metal wire or a metal plate, at a constant rate which is not influenced by the presence or absence of gas in the space surrounding the metal. The reader may think either of thermionic electrons flowing spontaneously out of a hot wire, or of photo-electrons flying out of a metal plate upon which ultra-violet light is shining.[1]

2. The Flow of Electrons Through a Very Rarefied Monatomic Gas, and Their Encounters with the Atoms

Conceive a source of electrons, a negative electrode or cathode, which is enclosed in a tube. If the tube is highly evacuated, the

[1] While forming one's ideas it is preferable to think of the photoelectric source, for a variety of reasons; the electron-stream is not very dense, the electrons emerge with kinetic energies never in excess of a certain sharply-marked limiting value, the metal is cold and not likely to react chemically with whatever gas surrounds it. Also several of the classical fundamental experiments were performed in the years from 1898 to 1906, when the photoelectric effect had become a reliable instrument of research and the thermionic effect had not. Nowadays it is sometimes used in the hope of surpassing the accuracy of earlier work, or in experiments on compound gases which the hot wire might decompose. Still the hot wire is so much easier to insert and handle, its emission so much more convenient and controllable, that it will no doubt be employed in the great majority of experiments in the future as in the past.

electrons enter the vacuum freely; electricity has no horror of a vacuum, any more than nature generally. Still there is something which suggests the *horror vacui* of the scientists before Galileo; for the electrons which are already partway across the vacuum tend, by their electrostatic repulsion, to push back their followers which are just emerging from the metal. This is the space-charge effect, which has become famous since the audion became almost as common an object as the incandescent lamp in the American home. I shall presently have to write down the equations describing this effect; for the time being we may ignore it, so long as the electron-stream is not more profuse than a photoelectric current generally is. The electrons of these scanty discharges enter into the vacuum and pass over without hindrance.

At this point it is advisable to say what is meant by a "vacuum." Scientists are growing more exigent year by year in their use of this term; thirty or forty years ago people spoke of "vacuum tubes" meaning tubes so full of gas that they would transmit a big current with a resplendent luminous display, but this self-contradicting usage has become quite intolerable. At the present day the least density of gas, or the highest vacuum, commonly attained corresponds to a gas-pressure about 10^{-11} as great as the pressure and density of the atmosphere. This means that there are about 10^{-8} molecules in a cubic centimetre of the "vacuum," which may make the name sound absurd. But the practical criterion for a vacuum is not whether the remaining atoms seem many or few, but whether they are numerous enough to affect the passage of a discharge; and as an electron shooting across a tube 10 cm. wide and evacuated to this degree has 999999 chances out of a million of getting clear across without encountering a molecule, the tube is vacuous enough for any sensible definition.

Next we will imagine that a gas is introduced into the tube, in quantity sufficient so that each electron going from cathode toward anode will collide on the average with one or possibly two atoms on its way. It is best to begin by thinking of one of the noble gases, of which helium, argon and neon are the ones in common use; or of the vapour of a metal, mercury vapour being much the easiest of these to work with; for their atoms behave in a simpler and clearer manner toward the electrons than do the molecules of the commonest gases, particularly the oxygen molecules which are so numerous in air. In fact the practice of using the noble gases and the metal vapours—that is to say, the *monatomic* gases—wherever possible in these researches ought really to be regarded as one of the great advances of the last few years; our predecessors would certainly have learned more about the discharge in gases than they ever did, if they had not studied it in air ninety times out of a hundred, and in other diatomic gases most of the other ten.

Let us suppose that the tube contains helium of the extremely small density I have just defined. Then so long as the kinetic energy of an electron does not exceed 19.75 volts, it will rebound from any helium atom which it strikes, like a very small perfectly elastic ball rebounding from a very large one. We might conceive the contents of the tube (for this purpose and only for this purpose!) as a flock of immense ivory pushballs floating languidly about, with a blizzard of equally elastic golfballs or marbles darting through the interspaces and occasionally striking and bouncing off from one of the pushballs. If the collisions between electrons and atoms are perfectly elastic, as I have said without giving evidence, the electron will lose an extremely small part of its kinetic energy at each collision, owing to the great disparity in masses—a fraction varying from zero up to not more than .000537 depending on the direction of rebound.

This was verified in a pretty experiment by K. T. Compton and J. M. Benade, who utilized a certain effect[2] which electrons produce when they have kinetic energy exceeding 19.75 volts at the moment of a collision with a helium atom. For example, when the pressure of helium was 4.34 mm. and the electrons were drawn from a cathode to an anode 0.265 cm. away, a voltage-difference of 20.25 (plus an unknown correction) was required to produce this effect; when the anode was 0.90 cm. from the cathode the required voltage-difference was 23.45 (plus the same correction). The extra volts were spent in replacing the energy lost by the electrons in the collisions with helium atoms over the extra 6.3 mm.; they amounted to an average of .0003 of the electron's energy lost in each collision, excellently in agreement with the assumption.

Now as for the transit of the electron-stream from cathode to anode, the helium atoms will simply thin it down by intercepting some of the electrons and turning their courses backwards or aside. The greater the number of atoms in the path, the greater the proportion of electrons intercepted; it can easily be seen that, so long as the gas is not denser than I have specified, this proportion increases as an exponential function of the number of atoms between cathode and anode,[3] whether this number be increased by introducing more gas or by moving the anode farther away from the cathode. If

[2] Incipient ionization, as described below.

[3] The proportion increases more slowly when there are already so many atoms between anode and cathode that an electron is likely to strike two or more on its way across.

the anode and the cathode are two parallel plates d centimetres apart, and there are P helium atoms in a cubic centimetre of the gas between, and N_0 electrons start out in a second directly towards the anode from any area of the cathode, the proportion $\Delta N/N_0$ of electrons which are intercepted before they reach the anode is

$$\Delta N/N_0 = 1 - e^{-APd} \tag{1}$$

and the number of electrons reaching the corresponding area on the anode in a second, $N_0 - \Delta N$, conforms to the equation:

$$\log_\epsilon (N_0 - \Delta N) = -APd + \text{const.} \tag{2}$$

The coefficient A is a constant to be interpreted as the effective cross-sectional area of the helium atom relatively to an oncoming electron—that is, the atom behaves towards the electron like an obstacle presenting the impenetrable area A to it.

In the experiments performed to verify these assertions and determine the value of A, the simple geometrical arrangement which I have described is generally modified in one way or another for greater accuracy or convenience. Mayer approached most nearly to the simple arrangement; in his apparatus (Fig. 1) the electrons

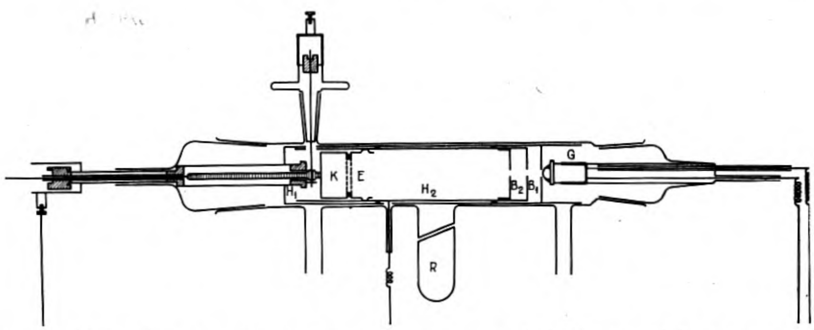


Fig. 1—Apparatus for determining the percentage of electrons which go across a gas of variable thickness without interception. (Mayer, *Annalen der Physik*)

which emerge from the hot filament at G, pass through the two slits in front of it, and then go down the long tube to the anode K, which is drawn backward step by step. The logarithmic curves of current versus distance for various pressures of nitrogen (Fig. 2) are straight. Unfortunately the current also diminishes as the distance is increased when the nitrogen is pumped out altogether; this is attributed partly to residual vapors and partly to the electrons striking the walls of the tube. The other curves are corrected for this effect, and then A is calculated. For helium it is 25.10^{-16} cm^2; the values obtained by modifications of the method agree well.[4]

The helium atoms therefore behave as so many minute and yet appreciable obstacles to the passage of the electron-stream, so long as the electrons are not moving so rapidly that their energies of motion do not surpass 19.75 volts. Electrons as slow as these bounce off from the atoms which they strike. When, however, an electron possessing kinetic energy greater than 19.75 volts strikes a helium atom,

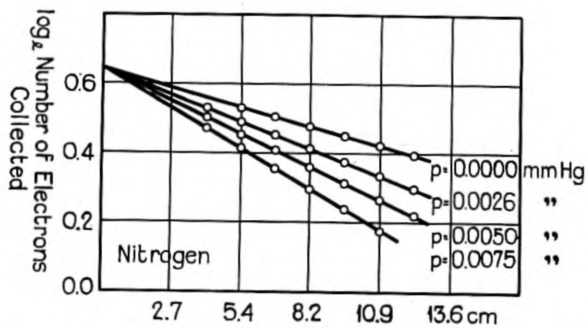


Fig. 2—Curves illustrating the interception of electrons by nitrogen molecules which they strike. (Mayer, *Annalen der Physik*)

it is liable to lose 19.75 volts of its energy to the atom, retaining only the remainder. This energy does not become kinetic energy of the atom, a process which would be incompatible with conservation of momentum; neither is the atom broken up; it receives the quota of energy into its internal economy, where some kind of a domestic change occurs with which we are not concerned for the moment, except in that it furnishes an exceedingly accurate indirect way of calculating the exact amount of energy taken from the electron. The atom is said to be put into an "excited" or sometimes into a "meta-

[4] The modified methods are generally more accurate. Ramsauer's device, which I described in the first article of this series, is probably the best. By a magnetic field he swung a stream of electrons around through a narrow curving channel, and those which were deviated even through a few degrees struck the limiting partitions and were lost from the beam; he varied the number of atoms in the channel by varying the gas-pressure. In this way he discovered that A for argon atoms differs very greatly for different speeds of the electrons; it was later found that other kinds of atoms have a variable A, although happily the variations are not great. This seems strange at first, but it is probably stranger that A should have nearly the same value for different speeds of the oncoming electrons, as for many atoms it does; and stranger yet that it should have the same value for an oncoming atom as for an oncoming electron, as is often tacitly assumed, and not too incorrectly.

stable" state, and the energy which it takes up, measured in volts, is called its *resonance-potential*. The electron is left with only the difference between its initial energy and the 19.75 volts which it surrendered.

This loss of energy in a so-called "inelastic" collision can be dem-

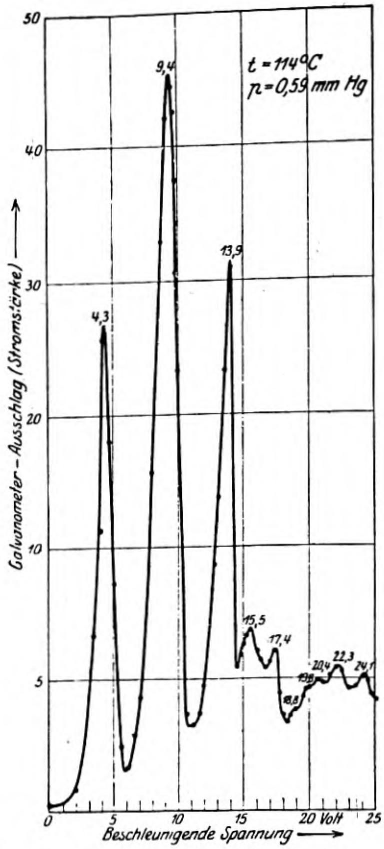


Fig. 3—Curve displaying resonance-potentials of mercury. (Einsporn, *ZS. f. Physik*)

onstrated by inserting a third electrode into the path of the electrons, charged negatively to just such a degree that an electron retaining its full initial speed can overcome the repulsion of the electrode and win through to it, while an electron which has lost a quantity of its kinetic energy in an inelastic collision cannot quite "make the grade." When the energy of the electrons streaming into the helium is raised just past 19.75 volts there is a sudden falling-off in the number of electrons arriving at the third electrode. The curve in Fig. 3, obtained by Einsporn, shows the current into such an electrode in mercury-vapor rising and falling again and again as the voltage passes through the values which are integer multiples of 4.9 volts, the least resonance-potential of mercury. Helium has a second resonance-potential, at 20.45 volts; neon has two, at 16.65 and 18.45 volts respectively; argon three, at 11.55, 13.0 and 14.0 volts;[5] mercury two, at 4.9 and 6.7 volts. It is almost certain that in each case these are only the most conspicuous among many, but the lowest mentioned is the lowest of all.

Up to this point we find the gas acting as a mere inert obstruction to the discharge; every collision of an electron with an atom interrupts the progress of the electron toward the anode and to that extent impedes the discharge. Past the resonance-potential the same action continues, although the interruption is doubtless less severe when the electron is deprived of part of its energy of forward motion than when it is flung backward with its motion reversed in direction and its energy intact. At the resonance-potential, the gas does begin to assist the discharge in an indirect way. Atoms which are put into an "excited state" by a blow from an electron revert of themselves to the normal state, some time later; in so doing they emit radiation, some of which falls upon the cathode; some of this is absorbed in the cathode metal, and expels electrons which go along with the maintained electron-stream as extra members of it. Thus the gas helps in increasing and maintaining the discharge; this effect is of great theoretical importance, and I will return to it later; but in these actual circumstances it is not very prominent.

The really powerful cooperation of the gas in the discharge commences when the electrons are given so great an energy that they disrupt the atoms which they strike, tearing off an electron from each and leaving a positively-charged residue, an *ion* which wanders back towards the cathode while the newly-freed electron and its liberator go on ahead towards the anode. The onset of this ionization may be detected by inserting a third electrode into the gas, it being charged negatively to such a degree that no electrons can reach it, but only positive ions; or by the increase in the current between cathode and anode, for the current increases very suddenly and very rapidly when the energy of the primary electrons is raised past the threshold-value, the *ionizing-potential* of the gas; 24.5 volts for helium, 21.5 for neon, 15.3 for argon, 10.4 for mercury. Consider for example the

[5] I take the values for neon and argon from Hertz' latest publication.

precipitate upward rush of the current-voltage curve in Fig. 4, from the work of Davis and Goucher.[6]

At this point I will digress to speak very briefly of the succession of events which occurs when the electron-stream is much denser than

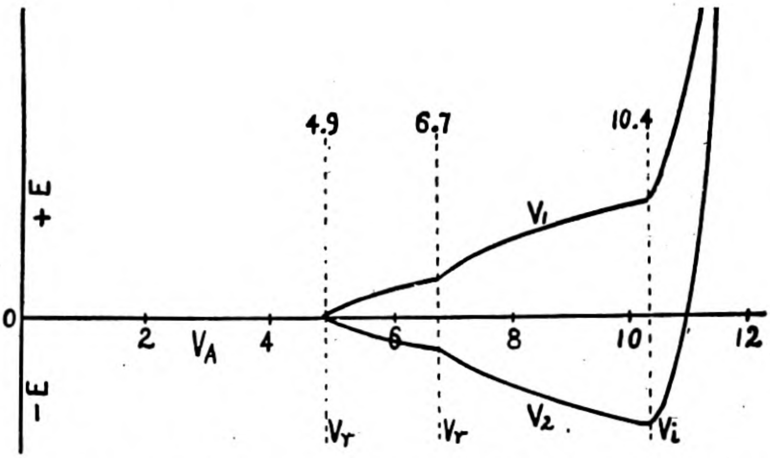


Fig. 4—Onset of ionization in mercury vapor at 10.4 volts (preceded by subsidiary effects at 4.9 volts and 6.7 volts; see footnote[6]). (Davis and Goucher)

we have hitherto imagined. So long as the energy of the electrons does not attain the resonance-potential of the gas, there is no reason to expect any novel effects; the collisions will be perfectly elastic, just as when the electrons were few. But when the atoms are thrown into the "excited state" by impacts, there will be occasional cases of an atom being struck twice by electrons in such quick succession that at the moment of the second blow, it is still in the excited state provoked by the first. Now, much less energy is required to ionize an atom when it is in the excited state than when it is normal; consequently when the electrons are so abundant that these pairs of

[6] The sudden upturn at 10.4 volts is the swift rise of current at the onset of ionization. The much less violent upturns at 4.9 and 6.7 volts are due to the electrons expelled from the metal parts of the apparatus by the radiation from the excited atoms. In the lower curve, by modifying the apparatus, the latter upturns are translated into downturns to distinguish them from the upturn which denotes ionization. This distinction was not realized until 1917, and in articles published between 1913 and 1917 the lowest resonance-potentials of gases are given as their ionizing potentials. Enormous improvements in the methods and technique of measuring these critical potentials, and recognizing of which kind they are, have been effected since then.

nearly-simultaneous collisions happen often, ionization will begin at the resonance-potential. *In a profuse electron-stream, the threshold potential for ionization is the lowest resonance-potential.* Another feature of the profuse discharge is, that when ionization does commence the current leaps up much more suddenly and violently than it does in the scanty discharge. This is because the electron-current is depressed at first by the space-charge effect, the repellence which the electrons crossing the gap exert against the electrons which are on the verge of starting; when positive ions first appear in the gap, they cancel the action of a great number of the traversing electrons, and the flow of electrons from the cathode to anode is immensely increased. I shall speak of this more extensively further on.

We return to the case of the feeble electron-stream. We have considered various things which an electron may do to a helium atom which it strikes—bouncing off harmlessly, or putting the atom into an excited state, or ionizing it; we have mentioned that each of the two latter actions commences at a critical value of energy, at the so-called *resonance* or *ionizing* potential, respectively; we have considered the effect of each of these actions upon the discharge. Have we listed all the possible interactions between atoms of matter and atoms of electricity, when electrons flow across helium? and if we knew all the resonance potentials and all the ionizing potentials [7] of helium, could we predict all the features of all electrical discharges in pure helium, whether in rarefied gas or in dense, whether the electron-stream be scanty or profuse? This is the general belief; whether justified, it is impossible to say. We evidently need another Maxwell or another Boltzmann, somebody exceedingly skilful in statistical reasoning, able to take the information we can provide about the possibility or the probability of various kinds of impacts, and deduce the state of affairs in the mixture of atoms, ions and electrons without getting hopelessly entangled in the frightful maze of equations into which his very first steps would certainly lead him. While awaiting him we have to content ourselves with our successes in interpreting the flow of electrons through very rarefied helium and the other noble gases and the metal vapors; and as for the discharges in denser gases

[7] I have simplified this passage somewhat so as not to retard the exposition. We know that an electron may "excite" a helium atom if its energy exceeds 19.75 volts, but this does not prove that it *must* do so; it is more reasonable to suppose that it has a certain chance of exciting the atom, zero when its energy is less than 19.75 volts, but greater than zero, and a certain function of its energy, when the latter exceeds 19.75 volts. We should know these functions for all the resonance-potentials and for the ionizing-potential; independent experiments to determine them have been performed, and no doubt will be multiplied.

we have to take the experimental data as we find them, and analyze them as best we may, not with too great an expectation of penetrating to the properties of the ultimate atoms; and yet, as we shall see, the analysis does in certain cases penetrate unexpectedly far.

3. The Flow of Electrons Across Dense Air, Nitrogen, Hydrogen and Similar Gases

The celebrated series of researches by Professor Townsend of Oxford and by his pupils, commenced in 1902 and continuing through the present, relate chiefly to such gases as hydrogen, nitrogen, oxygen and the familiar mixture of the last two which we breathe; and chiefly to these gases at densities much greater than we have hitherto considered—densities corresponding to such pressures as a thousandth or a hundredth of an atmosphere, therefore so great that an electron crossing over from a cathode to an anode a few centimetres away must collide with scores or hundreds of atoms. If a stream of electrons is poured into perfectly pure helium of such a density, we must not look for a sudden onset of ionization when the voltage between cathode and anode is raised just past 24.5, for the reason illustrated by those experiments of Compton and Benade—the electrons lose energy in all of their collisions, even the elastic ones, and arrive at the anode not with the full energy corresponding to its potential but with this energy diminished by what they lost on the way. In the familiar diatomic gases, the electrons lose much more energy in their ordinary collisions. I did not speak of these gases in the foregoing section, because experiments of the very same type as those which show the sharp distinction between elastic impacts and inelastic impacts in the noble gases and give the sharply-defined values of the resonance-potentials of these gases, yield comparatively vague and ill-defined data, when they are performed on hydrogen or air. In these gases, above all in active gases like oxygen or iodine, it is unlikely that any of the impacts, whether the electrons be moving rapidly or slowly, are truly elastic.[8]

[8] However, Foote and Mohler have obtained quite undeniable evidence of critical potentials, at which the loss of energy by the impinging electron is much greater than it is just below these potentials. The electron can transfer energy to (and receive energy from) a molecule in more different ways than to (from) an atom; such as by setting the molecule into rotation, or putting its constituent atoms into vibration relatively to one another. There is also the mysterious fact of "electron affinity"—an electron may adhere firmly to a non-ionized molecule. Numerous measurements of the rate at which electrons progress through a gas (a field of research which I have not space to consider here) indicate that at field strengths such as prevail in these experiments, adhesion of electrons to molecules is rare and transient.

Now if an electron on its way through the electric field from cathode to anode strikes atoms so often that it rarely has a chance to acquire more than say half a volt of energy from the field between one impact and the next, and if in each impact it loses most of the energy it has just acquired—if this condition prevails, we need not wonder that the voltage between the electrodes must be raised far beyond the ionizing-potential of the gas before there is the least sign of intensification of current.

In interpreting the experiments upon such gases and at such pressures as these last, it has been customary to make a more drastic assumption, the opposite extreme from the one which justified itself in dealing with rarefied helium; it is assumed that the electron surrenders at every impact all the energy which it has derived from the field since its last preceding impact. One may be inclined to make mental reservations in accepting so extreme an assumption, and it could almost certainly be advantageously modified; but as a tentative assumption it is successful enough to be legitimate. If it is true the electron can never build up a capital of energy step by step along its path; the only chances it will have to ionize will come at the ends of unusually long free flights.

Let us imagine a specific case *pour fixer les idées*: supposing the anode and the cathode to be parallel plates d apart, and representing the potential-difference between them by V and the field strength between them by X ($X = V/d$), we will set $d = 6$ cm., $V = 300$ volts, $X = 50$ volts/cm.; we will imagine that the interspace is filled with a gas having an ionizing-potential equal to 15 volts, and so dense that the average free path of an electron between collisions is one millimetre. I say that the *average* free path is 1 mm. long; if all the sixty free paths which the electron traverses in going from cathode to anode were equal, it would never acquire more than 5 volts of energy, and could never ionize an atom; but owing to the statistical distribution of free paths about the mean value, there will be a certain number out of the sixty which will be longer than three millimetres, and long enough, therefore, for the electron to acquire the 15 volts of energy which are necessary to ionize. In this case there will be $60/\epsilon^2$, about eight, of these long free paths. In each centimetre there will be $10/\epsilon^2$ of them. I will use the letter α' to designate this latter number, which is the number of atoms struck by the electron in each centimetre of its path, at moments at which it has energy enough to ionize an atom; α' is therefore the *number of chances to ionize* which the electron has *per centimetre*. The formula for α' is:

$$\alpha' = \frac{1}{\lambda}\epsilon^{-V_0/X\lambda} = C\epsilon^{-CV_0/X} = Bp\epsilon^{-BpV_0/X}, \tag{3}$$

in which V_0 represents the ionizing-potential of the gas; λ represents the mean free path of the electron; C, its reciprocal, is the number of collisions suffered by the electron in each centimetre of the path; and, since C is proportional to the pressure of the gas, it is replaced by Bp in the final formulation.[9]

It is already clear that the new assumption leads to a theory which requires a different language and a different set of ideas from those of the foregoing section. Not the ionizing-potential, but the number of ionizations performed by an electron in a centimetre of its path, is the quantity to be measured by experimental devices; not the voltage between the electrodes, but the field strength in the gas, is the factor which controls the phenomena.[10] In dealing with gases which are expected to conform to the theory, the appropriate procedure is to measure the number of molecules which an electron ionizes in a centimetre of its path, for all practical values of the field strength X and the density of the gas (or its pressure p) as independent variables. I will designate this number, following the usual practice, by α; if the theory is true it cannot be greater than α', it may be less. These quantities α and α' are statistical quantities, not like the ionizing-potential qualities of the individual atom or molecule, and this is a misfortune and disadvantage of the theory and of the experiments which it interprets; we are not, so to speak, in the presence of the ultimate atoms as before, we are one step removed from them, and this step a difficult one to take.

The measurement of α is effected by varying the distance d between anode and cathode, and determining the current as function of d. If N_0 electrons flow out of the cathode in a second, the ionization commences at the distance $d_0 = V/X$ from the cathode, and from that

[9] Since the number of free paths, out of a total number N_0, which exceed L in length is equal to $N_0 \exp(-L/\lambda)$; and since the potential-difference between the beginning and the end of the path of length L, if parallel to the field, is XL. It may be objected that the electrons bounce in all directions from their impacts, while the language of this paragraph implies that they are always moving exactly in the direction of the field. The rebuttal is, that if they do lose almost all of their energy in an impact, or all but an amount not much greater than the mean speed of thermal agitation, they will soon be swerved around completely into the direction of the field no matter in what direction they start out.

[10] The ionizing-potential determines the distance from the cathode at which ionization commences; this is equal to $d_0 = V_0/X$, and within this distance from the cathode there is no ionization and the theory does not apply; beyond this distance the ionization is controlled entirely by the field strength and by the number of inflowing electrons and the voltage between cathode and anode affects it only insofar as it affects these.

point onward the electron-stream increases exponentially, so that the current Ne arriving at the anode is

$$Ne = N_0 e \exp \alpha \, (d - d_0) \tag{4}$$

In Townsend's experiments the cathode was a zinc plate, the anode a film of silver spread upon a quartz plate; through little windows in the silver film a beam of ultraviolet light entered in from behind, crossed over the interspace and fell normally upon the zinc plate, and drove electrons out of it. The zinc plate was raised and lowered by a screw; the voltage-difference between it and the silver film was altered *pari passu* so that the field strength in the gas remained always the same. The current rose exponentially as the distance between the plates was increased, and thus α was determined. A typical set of data (relating to air at 4 mm. pressure, with a field strength of 700 volts/cm.) is plotted logarithmically in Fig. 5, the logarithm of the current as ordinate and the distance from anode to cathode as abscissa. The first few points lie close to a straight line, corresponding to an exponential curve such as equation (4) requires; the value deduced for α is 8.16. (The distance d_0 is about .35 mm. and has been ignored.) Of the divergence of the later points from the straight line I will speak further on.

Such an experiment shows that there *is* an α—that the theory is not at any rate in discord with the first obvious physical facts—and it gives the value of α for the existing values of X and p. Townsend performed many such measurements with different field strengths and different pressures, and so accumulated a large experimental material for determining α as function of the two variables p and X. To interpret these we will begin by making the tentative and temporary assumption that whenever a molecule is struck by an electron having energy enough to ionize it, it *is* ionized—that is, $\alpha' = \alpha$. Rewriting the equation (3) which expresses α' as function of p and X, we see that

$$\alpha'/p = B \exp \left(-B V_0 p / X\right) = f \, (X/p). \tag{5}$$

Therefore, if $\alpha' = \alpha$, the quotient of α by p is a function of X and p only in the combination X/p; or, whenever the pressure and the field strength are varied in the same proportion, the number of molecules ionized by an electron in a centimetre of its path varies proportionally with the pressure. I leave it to the reader to invent other ways of expressing (5) in words which illuminate various aspects of its physical meaning.

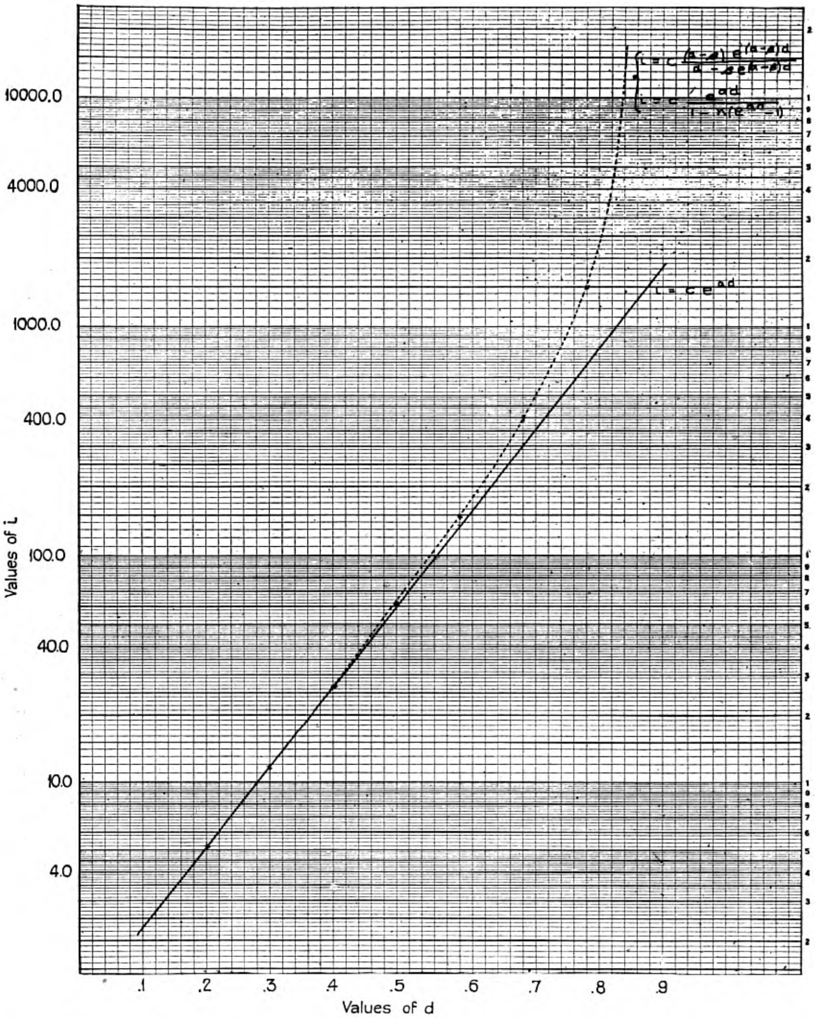


Fig. 5—Logarithmic plot of the currents across a gas (air) in which ionization by collision is occurring, for a constant field strength and various thicknesses of gas (Data from Townsend)

Experimentally, the test of (5) is made by dividing each one of Townsend's values of α by the pressure at which it was determined, and then plotting all these values of α/p versus the corresponding values of X/p. All the points for any one gas should lie on or close to a single curve, and within certain ranges of pressure and field strength they do; so far, good. The curve should be an exponential one, and within certain ranges of field strength and pressure it is; again, good. The next step is to calculate the values of B and V_0 which the curve imposes on the gas to which it relates. I quote the values of V_0, the ionizing-potential, which Townsend presents:

Air	N_2	H_2	CO_2	HCl	H_2O	A	He
25	27.6	26	23.3	16.5	22.4	17.3	12.3

When the first of these values were determined, no more direct way of measuring ionizing-potentials was known. Now that we have some values obtained by the direct methods sketched a few pages back, and fortified by indirect but very forcible evidence from spectroscopy, it is possible and quite important to test some of these. The values for argon and helium, although of the proper order of magnitude, are certainly too low. This is not in the least surprising, considering how many of the collisions between electrons and atoms must be perfectly elastic. It seems indeed rather mysterious that the current-voltage relation in either of these gases should have conformed closely enough to (4) to make it possible to define and measure α; but the electrons no doubt entered into many of the collisions with energy enough to put the atoms into excited states, if not to ionize them; and it is nearly always possible to take refuge in the assertion that the impurities may have been sufficient to distort the phenomena. As for the other gases in the list, all of them diatomic or triatomic, Townsend's values are too high—not very much too high, however; usually a matter of one-third to two-thirds.[11]

It appears therefore that the theory I have just developed is too simple, and must be amended. It seems natural to begin by dropping the tentative assumption that a molecule is ionized whenever it is hit by an electron having as much or more energy than is required to ionize it, and adopt instead the idea once already suggested in these pages, that it is sometimes but not always ionized by such a blow; that there is a certain *probability* of ionization by a blow from an electron having energy U, a probability which is zero when $U < V$ and is some yet-to-be-determined function of U when $U > V$. This would leave intact the conclusion that α/p should be a function of X/p, a conclusion which we have already found to be verified by experiment; but it would relieve us of the necessity of assuming that

[11] Townsend's values of B likewise correspond to values of the effective cross-section of the molecule, the quantity A of equation (2), which are of the same order of magnitude as the directly determined values of A.

that function is precisely the exponential function appearing in (5). Essentially the theory is reduced to this postulate: the number of molecules ionized by an electron in a centimetre of its path depends only upon the energy it acquires from the field in its free flight from one collision to the next. If in this form the theory still cannot give satisfaction, the next step will be to alter the original assumption that the electron comes practically to a dead stop in every collision. In dealing with the noble gases and the metal vapours, the facts about elastic collisions which I have already outlined prove that this assumption should not be made at all. It is clear that this is another problem for the future Boltzmann!

Meanwhile, one of the cardinal features of the Townsend experiments is the fact that they display the gradual advent of the transformation of the maintained currents which we have hitherto considered, into the self-maintaining discharges which are the familiar and the spectacular ones; and we now have to examine the agencies of this transformation.

4. The Discharge Begins to Contribute to the Electron-Stream Which Maintains It

Greatly though the current of primary electrons from the cathode to the anode may be amplified by the repeated ionizations which I have described, there is nothing in this process which suggests how the discharge may eventually be transformed into a self-maintaining one like the glow or the arc. The free electrons may ionize ever so abundantly, but as soon as the supply from the cathode is suspended by cutting off the heat or the light, the last electrons to be emitted will migrate off towards the anode, and whatever electrons they liberate will go along with them, leaving a stratum of gas devoid of electrons in their wake; and this stratum will widen outwards and keep on widening until it reaches the anode, and then the discharge will be ended. Something further must happen continually in the gas through which the electrons are flowing, something which continually supplies new free electrons to replace, not merely to supplement, the old ones which are absorbed into the anode and vanish from the scene.

We have already noticed one sort of event continually happening in such a gas as helium traversed by not-too-slow electrons, which might conceivably develop into a mechanism for maintaining the discharge; for, when an atom of the gas is put into the "excited state" by a blow from an electron, it later returns into its normal state, and in so returning it emits a quantum of radiant energy which may strike the cathode, and be absorbed by it, and cause another electron to leap out of the cathode and follow the first one. There are two other conceivable processes, which have the merit that they can not only be conceived but also witnessed in operation by themselves when the right conditions are provided. Positive ions flung violently against a metal plate drive electrons out of it, as can be shown by putting a positively-charged collector near the bombarded plate and noticing the current of negative charge which flows into it; and positive ions flowing rapidly across a gas ionize some of the atoms in it, as may be shown by sending a beam of such ions across the interspace between two metal plates, with a gentle crosswise field between them which sucks the freed electrons into the positive plate. The mechanism of the first process is not understood, except when the positive ions are so many and so swift that they make the metal hot enough to emit thermionic electrons, which does not happen in the cases we are now considering. The mechanism of the second process is only dimly understood, but it is clear enough that a positive ion driven against an atom is much less likely to ionize it, than an electron of equal energy would be.[12] Either of these two processes is very inefficient, at least at the comparatively low speeds with which positive ions move under the circumstances of these experiments; but they are probably efficient enough to do what is required of them. No doubt all three of them contribute to the discharge; but the relative proportions in which they act certainly differ very much from one sort of discharge to another, and will furnish research problems for years to come.

Returning to Fig. 5, we note once more that as the electrodes are moved farther and farther apart while the density of the gas and the field strength are held constant, the current at first rises exponentially (linearly in the logarithmic plot) as it should if the free electrons and only the free electrons ionize; but eventually it rises more rapidly and seems to be headed for an uncontrollable upward sweep. Townsend attributed this uprush to the tardy but potent participation of the positive ions, either ionizing the molecules of the gas by impact after the fashion of the negative ions, or driving electrons out of the cathode when they strike it, or both. Either assumption leads to

[12] If momentum is conserved in the impact between ion and atom, the ion must retain a large part of its kinetic energy after the collision, or else the struck atom must take a large part of it as kinetic energy of its own motion; it is not possible for the striking particle to spend nearly its entire energy merely in liberating an electron from the struck one. Conservation of momentum perhaps does not prevail on the atomic scale; but of all the principles of classical dynamics, it is the one which the reformers of physics most hesitate to lay violent hands upon.

an equation expressing the data equally well. If we adopt the former, and designate by β the number of molecules ionized by a positive ion in a centimetre of its path, and by N_0 the number of electrons supplied per second at the cathode, we get

$$N=\frac{N_0(\alpha-\beta)e^{(\alpha-\beta)d}}{\alpha-\beta e^{(\alpha-\beta)d}}. \tag{6}$$

Of course, β must be much smaller than α, or the positive ions would have made themselves felt earlier. Or if we adopt the latter idea, and designate by k the number of electrons expelled from the cathode (on the average) by each positive ion striking it, we arrive at the formula

$$N=\frac{N_0e^{\alpha d}}{1-k(e^{\alpha d}-1)}. \tag{7}$$

Naturally k must be much smaller than unity for the same reason. In Fig. 5 the broken curve represents (6), with the values 8.16 and .0067 assigned to α and β; it also represents (7), with the values 8.16 and .00082 assigned to α and k.[13] (It was expected that the curves representing the two equations would be perceptibly apart on the scale of Fig. 5; but they were found to fall indistinguishably together.)

Evidently, therefore, the positive ions, weak and lethargic as they are in liberating electrons (one has only to compare β with α, or look at the value assigned to k in the last sentence!), can produce a notable addition to the current when the electrodes are far enough apart; and more than a notable addition, for when the distance d is raised to the value which makes the denominator of (6)—or of (7), whichever equation we are using—equal to zero, the value of N is infinite! Per-

[13] The derivations of (6) and (7) are as follows. Represent by $M(x)$ the number of electrons crossing the plane at x in unit time (the cathode being at $x=0$ and the anode at $x=d$); by $P(x)$ the number of positive ions crossing the plane at x in unit time; by N_0, the number of electrons *independently* supplied at the cathode per unit time, which is not necessarily equal to the value of M at $x=0$ (hence the notation); by i the current, or rather the current-density, as all these reasonings refer to a current-flow across unit area. We have

$$Me+Pe=i, \text{ hence}$$

making the assumption which leads to (6) we have

$$dM/dx=\alpha M+\beta P=(\alpha-\beta)M+\beta i/e$$

The boundary conditions are: $M=N_0$ at $x=o$ and $M=i/e$ at $x=d$. Integrating the equation and inserting these we get (6). Making the assumption which leads to (7) we have

$$dM/dx=\alpha M$$

The boundary conditions are: $M=N_0+k(i/e-M)$ or $(1+k)M=N_0+ki/e$ at $x=0$, and $M=i/e$ at $x=d$. Integrating the equation and inserting these we arrive at (7).

haps the best way to conceive of this is, that as the distance between the plates is increased toward that critical value of d, the value of N_0—which is the rate at which we have to supply electrons at the cathode, in order to keep a preassigned current flowing—diminishes continuously and approaches zero; so that eventually the current will keep itself going (and actually start itself) with the assistance of the occasional ions which are always appearing spontaneously in every gas, even though it be encased in an armor-plated shield. Of course, it is rather risky to predict just what is going to happen, when an equation which has been fixed up to represent a finite physical phenomenon over a certain range exhibits an infinite discontinuity at a point outside of that range. Usually, of course, the infinite value which the equation requires is modified into a finite one by the influence of some factor which was neglected when the equation was devised. In this case, however, the infinite discontinuity corresponds to a sudden catastrophic change. If an electrometer is shunted across the interspace between anode and cathode, its needle is forcibly jerked; if a telephone-receiver is connected in series with the interspace, it makes a clicking or a banging sound; if the gap is wide, so that the voltage just before the disruption is high, there is a brilliant flash, which may bear an uncomfortably strong resemblance to the lightning-bolt which is the cosmical prototype of all electric sparks.

What goes on after the critical moment of transition or transformation depends on many things; and not only on obvious features of the spark-gap, such as the kind and density of gas and the shape and size and material of the electrodes, but also on such things as the resistances and the inductances in series with the discharge, and the qualities of the source of electromotive force and its ability to satisfy the demands for current and voltage which the new discharge may make. Sometimes these demands are too extravagant for most laboratory sources or perhaps for any source to meet; probably this is why the spark between extended plane surfaces in dense air is as ephemeral as it is violent. But this does not always happen; in a sufficiently rarefied gas, the self-maintaining discharge which sets in after the transformation requires only a modest current and a practicable voltage, and supports itself with a few thousand volts applied across its terminals. The same thing occurs in a dense gas, if either of the electrodes is pointed or sharply curved, like a needle or a wire; the condition, more exactly, is that the radius of curvature of either electrode should be distinctly less than the least distance between the two. The transformation, however, is always very sudden, whether the new discharge be transient or permanent; and there are also sudden transitions from one sort of self-maintaining discharge to another, *e.g.*, from glow to arc or from one kind of glow to another when certain critical conditions are transgressed (critical conditions which may themselves depend on the battery and the circuit as well as the constants of the spark-gap). There are discontinuities of current and discontinuities of voltage at these transitions, and abrupt changes in the visible appearance in the discharge; and at each transformation there is a rearrangement of the distribution of space-charge in the gas. Hitherto we have encountered space-charge only in one or two of its simplest manifestations, retarding the flow of an electron-stream across a vacuum, and suddenly annulled when positive ions are mingled with the stream. Now we have to consider much subtler and more complicated cases, in which the space-charge varies rapidly in density and even in sign from one part of the gas to another, and the field and potential distributions are utterly distorted by it; and these distortions are essential to the life of the discharge. This distribution of space-charge is indeed dominant; and so I will write down some formulae which may be used to describe it.

5. Digression to Write Down Some Space-Charge Equations

The fundamental equation of the electrostatic field, known as Poisson's equation, is

$$\nabla^2 V = \frac{d^2V}{dx^2} + \frac{d^2V}{dy^2} + \frac{d^2V}{dz^2} = -4\pi\rho \tag{8}$$

in which V represents the electrostatic potential, and ρ the volume-density of electric charge.

We consider only the mathematically simplest case in which all variables are constant over each plane perpendicular to the x-axis, and so depend only on the coordinate x; as for example near the middle of an exceedingly wide tube with the x-axis lying along its axis. In this case Poisson's equation is

$$\frac{d^2V}{dx^2} = \frac{dX}{dx} = -4\pi\rho \tag{9}$$

in which X represents the potential-gradient, or field strength with sign reversed.[14] The value of X is determined at all points when the

[14] Field-gradient is therefore, proportional to space-charge with sign reversed, and *vice versa*. Positive field-gradient implies negative space-charge; negative field-

value of X at any one point and the values of ρ at all intermediate points are preassigned. Thus let X_0 represent the preassigned value of X at $x=0$, and X_d represent the value of X at $x=d$; we have

$$X_d = -4\pi \int_0^d \rho \, dx + X_0. \tag{10}$$

Consequently the P.D. between any two points is also determined; that between $x=0$ and $x=d$ is

$$V_d - V_0 = -4\pi \int_0^d dx \int_0^x \rho \, dx + X_0 d. \tag{11}$$

Now we introduce the further assumption that the electric charge is concentrated upon corpuscles (electrons or charged atoms) of one kind, of equal charge E and mass m, of which there are ndv in a very small volume dv at x; n is a function of x. Then

$$nE = \rho. \tag{12}$$

Assume finally that the corpuscles are moving with speed u, identical for all corpuscles having the same x-coordinate, but depending on x; represent the current-density by i; we have

$$nEu = i \tag{13}$$

and consequently

$$\rho = i/u. \tag{14}$$

Now consider the flow of current between two parallel planes, from one electrode at $x=0$ to the other at $x=d$. If the current is borne by corpuscles of one kind, and the assumption last made is true; and if we know the speed of the corpuscles at every point between the plates, and the field strength at some one point; then we can calculate the field strength everywhere between the plates, and the potential-difference between them.

The customary convention about the field strength is to assume it to be zero at the electrode from which the corpuscles start, so that $X_0=0$ in (11). Rewriting (11) to take account of (14), we have

$$V_d - V_0 = -4\pi i \int_0^d dx \int_0^x dx/u \tag{15}$$

as the general equation.

gradient implies positive space-charge; uniform field implies zero space-charge. It is instructive to examine mappings of field-distribution with this principle in mind; such mappings, for example, as those in Fig. 9. The uniform field in a current-carrying wire means that positive and negative charges are distributed everywhere in the metal with equal density—a conclusion one might forget, but for these more general cases.

If we suppose that the corpuscles acquire their speed u at the distance x in free flight from the electrode where they start, we have $\frac{1}{2}mu^2 = eV$, and

$$(V_d - V_0)^{3/2} = \frac{9\pi}{\sqrt{2}}\sqrt{\frac{m}{E}}\, id^2. \tag{16}$$

This is the equation adapted to electrons or other ions flowing across otherwise empty space.

If we suppose that the corpuscles have at each point a speed proportional to the field strength at that point, we have $u = \pm k\, dV/dx$, and

$$V_d - V_0 = \frac{2}{3}\sqrt{\frac{8\pi i\, d^3}{k}}. \tag{17}$$

This equation would be adapted to ions drifting in so dense a gas, or so weak a field, that they acquire *very little* energy from the field (in comparison with their average energy of thermal agitation in the gas) between one collision and the next, and lose it all at the next.[15]

If we conceive of ions which acquire *much* energy from the field between one collision and the next (much, that is, in comparison with their average energy of thermal agitation) and lose it all at the next collision, we have $u^2 = (\pi el/2m)\, dV/dx$ and

$$(V_d - V_0)^{3/2} = Cid^{5/2} \tag{18}$$

the constant C being equal to $\sqrt{m/El}$ multiplied by a certain numerical factor, and l standing for the mean distance traveled by the ion between one collision and the next.

The theory just given is too simple; it is an essential fact of the actual physical case that the ions emerge, at the surface of the electrode whence they start, with forward velocities which are distributed in some way or other about a mean value. These initial forward velocities, though often small compared with the velocities which the ions may acquire as they cross to the other electrode, are large enough so that all of the ions would shoot across the gap if the field strength were really zero at the emitting electrode and assisted them everywhere beyond it. In fact the space-charge creates a retarding field at the surface of the emitting electrode, and a potential minimum (if the ions are negative; a potential maximum, if the ions are positive) at a certain distance in front of it. Here, and not at the emitting electrode as we previously assumed, the field strength is zero. Equation (16) is often valid in practice, because this locus of zero field-strength is often very close to the emitting electrode. In fact, by

[15] As in electrical conduction in solid metals (cf. my preceding article).

raising the P.D. between the plates sufficiently, the locus of zero field can be driven back into coincidence with the emitting plate; beyond which stage, the "limitation of current by space-charge" ceases. But if the P.D. is sufficiently low the potential minimum (or maximum) is prominent and is remote from the electrode, and in these cases the equations we have just deduced are inapplicable.

It thus may readily happen that when we apply a certain potential to one electrode and a certain other potential to another electrode separated from the first one by gas or vacuum, we may find points between them where the potential is *not intermediate* between the potentials of the electrodes. This is a queer conclusion, to anybody accustomed to the flow of electricity in wires. But it is true, and must be kept in mind.

6. The Self-Maintaining Discharges

The *Arc* ought to be the easiest to understand among the self-maintaining discharges, in one respect at least; for it keeps its own cathode so intensely hot that thermionic electrons are supplied continuously in great abundance at the negative end of the discharge, and the theorist can begin his labors by trying to explain how and why this high temperature is maintained. Anything which tends to lower the temperature of the cathode, for instance by draining heat away from it, is very perilous to the arc. Stark uses various schemes for preventing the cathode from growing very hot, and they all killed the arc. This also explains why the arc is most difficult to kindle and most inclined to flicker out when formed between electrodes of a metal which conducts heat exceptionally well, and most durable when formed between electrodes of carbon, which is a comparatively poor conductor for heat. It probably explains why the arc has a harder time to keep itself alive in hydrogen, a gas of high thermal conductivity, than in air. While the gas in which the arc has its being and the anode to which it extends both influence the discharge, the high temperature of the cathode is cardinal.

The cathode is presumably kept hot by the rain of positive ions upon it, striking it with violence and yielding up their energy of motion to it; at least this is the obvious and plausible explanation. Now the arc is commonly and easily maintained in fairly dense gases, with a comparatively small potential-difference between widely-separated electrodes; and the energy which an ion can acquire from the field strength prevailing in it, in the short interval between two collisions with molecules, is so small that it cannot be made to account for the furious heat developed at the cathode when the ions finally strike it. Just before the ions arrive at the cathode they must be endowed with a kinetic energy which is very unusual (to say the least) in the middle of the discharge; and it is in fact observed that just in front of the cathode there is a sharp and sudden potential-fall, corresponding to a strong field extending but a little way outward from the electrode and then dying down into the weak field prevailing through the rest of the arc. This strong field picks up the ions which have meandered to its outward edge from the body of the discharge and hurls them against the cathode—not very forcibly, for the energy they receive from that potential-fall is not a great amount by ordinary standards, and most of the ions probably lose some of it in collisions on the way; but with much more energy than they would be likely to possess anywhere else in the arc.

This potential-fall immediately in front of the negative electrode, the *cathode-fall* of the arc, is measured by thrusting a probe or sounding-wire into the discharge as close as possible to the cathode (generally about a millimetre away), and determining the *P.D.* between it and the cathode. The probe is regarded with some distrust, as it raises in an acute form the old question as to how far the phenomena we observe in nature are distorted by the fact that we are observing them; the wire may alter the potential of the point where it is placed, or it may assume a potential entirely different from that of the environing gas; but the general tendency nowadays, I believe, is to accept its potential as a moderately reliable index of the potential which would exist at the point where it stands if it were not there.[16] The cathode-fall, as so measured, depends unfortunately on quite a number of things; the material of the cathode, the gas, the current. The gas is always mixed with a vapour of the electrode-material, particularly in the vicinity of the electrode; the only way to have a single pure gas is to enclose the whole system in a tube, evacuate the tube to the highest possible degree, and then heat it until the vapor-tension of the metal of which the cathode is made rises high enough for the vapor to sustain the arc. This is practicable with the more fusible metals; and with mercury, the arc generates heat enough to maintain the vapor-tension sufficiently high. In pure mercury-

[16] On this matter the experiments of Langmuir and Schottky, mentioned further along in this article, promise new knowledge. The probe automatically assumes such a potential that the net current-flow into it is nil; for example, if it is immersed in an ionized gas in which electrons and ionized atoms are roaming about, its eventual potential is such that equal numbers of particles of the two kinds strike and are absorbed in it per unit time. If the electrons are much more numerous or have a much higher average energy, or both, this potential may be several volts more negative than the potential at the same point before the probe was put in. The same may be said about the wall of the tube.

vapor, the cathode-fall assumes the value 4.9 volts which is the first resonance-potential of the mercury atom and therefore, as we have seen, is effectively the ionizing-potential of the free mercury atom when the electron-stream is as dense as it is in the arc. This suggests a delightfully simple theory of the whole process: the electrons stream from the cathode, they acquire 4.9 volts of energy from the cathode-fall, they ionize mercury atoms at the outward edge of the region of high field strength, the positive ions so created fall backward across the cathode-fall and strike the cathode, surrender their energy to it and so keep it hot, more electrons pour out, and so forth *ad infinitum.* It remains to be seen whether so simple a theory can be modified, by statistical considerations or otherwise, to explain the values of the cathode-fall in mixed and diatomic gases.

We do not know *a priori* what is the ratio of the number of electrons flowing outward across the cathode-fall in a second to the number of ions flowing inward. It might, however, be very great, and still the number of ions within the region of the cathode-fall at any instant could far surpass the number of electrons within it—the electron moves so much more rapidly than the ion, and has so much better a chance of crossing the region in one free flight without a collision. Even in hydrogen, in which the ions are the lightest of all ions, the electron current would have to be 350 times as great as the ion-current if the electrons just balanced the ions in unit volume. It is therefore legitimate to try out the assumption that the region of cathode-fall is a region of purely positive space-charge, in which some such equation as (16), (17), or (18) gives the current of positive ions as a function of the cathode-fall and the width of the region. K. T. Compton selected (18). Unfortunately the width of the cathode-fall region has not been measured, but he assumed it equal to the mean free path of an electron in the gas. The value which he thus calculated for the current of positive ions was about 1% of the observed total current; the remaining 99% consists of the electrons.

From the cathode region onward to the anode, the gas traversed by the arc is dazzlingly brilliant. In the long cylindrical tubes which enclose the mercury arcs so commonly seen in laboratories and studios, the vapor shines everywhere except near the ends with a cold and rather ghastly white light tinged with bluish-green. This is the positive column of the mercury arc. The potential-gradient along it is uniform, suggesting the flow of electricity down a wire; but here the resemblance stops, for when the current-density goes up the potential-gradient goes down. The curve of voltage versus current, which for a solid metal is as we all know an upward-slanting straight line, is for the arc a downward-slanting curve (Fig. 6). Such a curve is called a *characteristic*, and the arc is said to have a *negative characteristic*. Ionization goes on continually within the positive column, and ions of both signs can be drawn out by a crosswise field; but recombination of ions, a process which we have not considered, also goes on continually and maintains an equilibrium. Presumably it

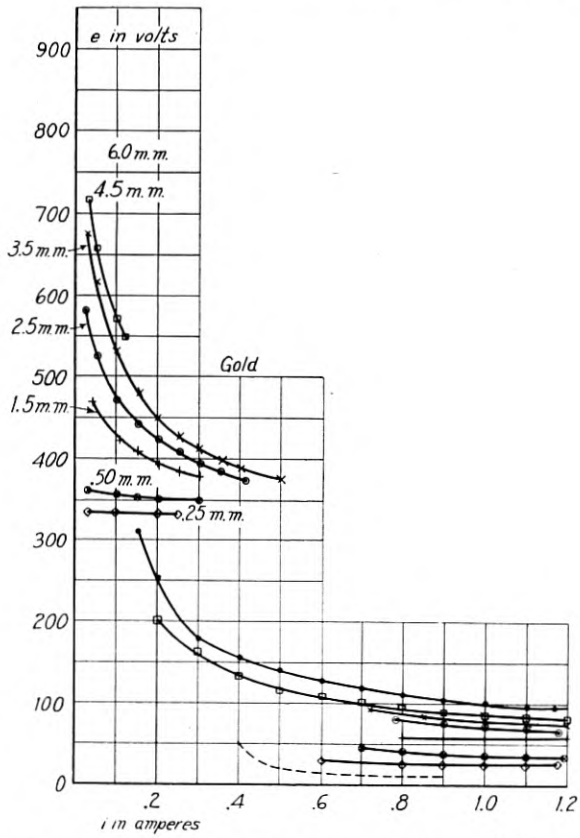


Fig. 6—Voltage-current curves or "characteristics," for arc discharges (below) and glow discharges (above) in air, between gold electrodes. The different curves correspond to different anode-cathode distances. (Ives, *Journal of the Franklin Institute*)

is the effect of the field strength on this equilibrium which causes the current-voltage curve to slant in what most people instinctively feel is the wrong way; but the theory of the equilibrium is not yet far advanced.

Langmuir and Schottky, working independently in Schenectady and in Germany, performed some very pretty experiments by thrusting negatively-charged wires or plates into the positive column. These wires and plates surrounded themselves with dark sheaths, the thickness of which increased as the potential of the metal was made more and more highly negative. The explanation is, that the electrons in the positive column cannot approach the intruded wire, being driven back by the adverse field; the dark sheath is the region from which they are excluded, and across it the positive ions advance to the wire through a field controlled by their space-charge. The equation selected by Langmuir to represent the relation between the thickness of the sheath, the voltage across it, and the current of positive ions into it, is (16). As the sheath is visible and its thickness can be measured, as well as the other quantities, the relation can be tested. This was done by Schottky; the result was satisfactory. When the intruded electrode is a wire, the sheath is cylindrical, and expands as the voltage of the wire is made more negative. As the area of the outer boundary of the sheath is increased by this expansion, more ions from the positive column touch it and are sucked in, and the density of flow of positive ions in the column can be determined. By lowering the potential of the wire gradually so that the electrons can reach it, first the fastest and then the slower ones, the velocity-distribution of the electrons in the column can be ascertained. Their average energy depends on the density of the mercury vapour, and may amount to several volts.

Beyond the positive column lies the anode, itself preceded by a sharp and sudden potential rise. The electrons are flung against it with some force, and it grows and remains very hot; usually, in fact, hotter than the cathode. This high temperature does not seem to be essential to the continuance of the discharge, for the anode can be cooled without killing the arc; yet it seems strange that a quality so regularly found should be without influence upon the discharge. One must beware of underestimating the influence of the anode; when an arc is formed in air between two electrodes of different materials, it behaves like an arc formed between two electrodes of the same material as the anode, *not* the cathode!

The so-called *low-voltage arc*, although not a self-maintaining discharge, merits at least a paragraph. A dense electron-stream poured into a monatomic gas from an independently-heated wire, and accelerated by a P.D. surpassing the resonance-potential of the gas, may ionize it so intensely that there is a sudden transformation into a luminous arc-like discharge. This is a sort of "assisted" arc, its cathode being kept warm for its benefit by outside agencies. Its history is a long and interesting chapter of contemporary physics, whereof the end is not yet. The most remarkable feature of this arc

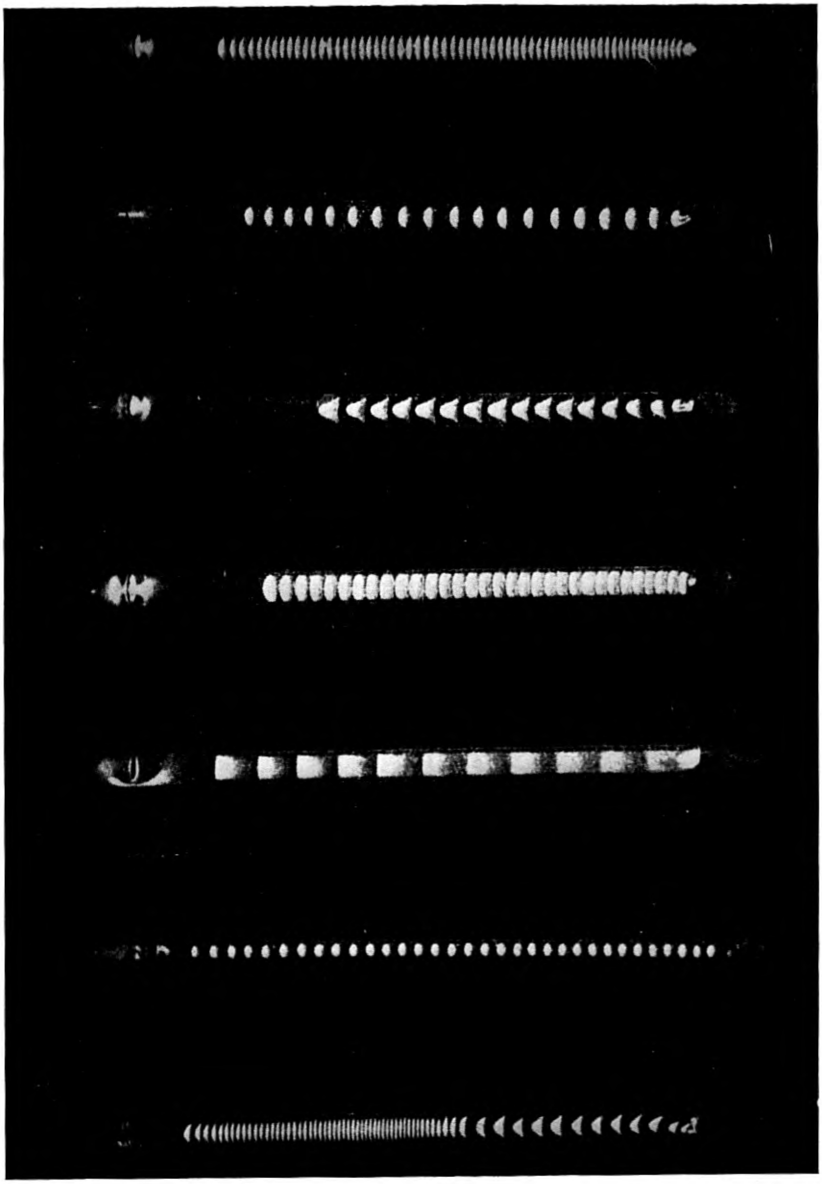


Fig. 7—Photographs of the glow-discharge in a long narrow cylinder, showing chiefly the subdivision of the positive column into striations. (De la Rue and Muller, *Philosophical Transactions of the Royal Society*)

is that it can survive even if the voltage between anode and cathode is far below the resonance-potential of the atoms of the gas, which seems impossible. A year ago it seemed that this effect could always be ascribed to high-voltage high-frequency oscillations generated in the arc. This explanation was presently confirmed in some cases and disqualified in others, and now it appears that when there are no oscillations an astonishingly strong potential-maximum develops within the ionized gas. Potential-maximum and oscillations alike are probably to be regarded as manifestations of space-charge.

The *Glow in a rarefied gas* is a magnificent sight when the gas is rarefied to the proper degree, not too little and not too much; divided into luminous clouds of diverse brightnesses and diverse colors, recalling Tennyson's "fluid haze of light," yet almost rigidly fixed in their distances and their proportions, it is one of the most theatrical spectacles in the repertoire of the physical laboratory. The grand divisions of the completely-developed discharge are four in number, two relatively dim and two bright; beginning from the cathode end,

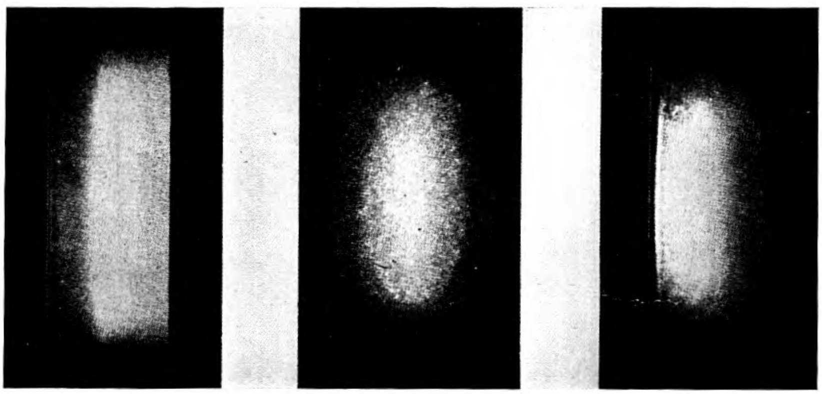


Fig. 8—The Crookes dark space between the cathode (thin line at left) and the negative glow. See footnote [12]. (Aston, *Proceedings of the Royal Society*)

they are the Crookes dark space, the negative glow, the Faraday dark space, and the positive column. Additional gradations of color and brightness can often be seen very close to the cathode and very close to the anode. Photographs of the glow which give anything approaching a true idea of its appearance to the eye are hard to find. I reproduce in Fig. 7 some photographs taken nearly fifty years ago by de la Rue, which have reappeared in many a text; they show chiefly the striking flocculent cloudlets into which the positive column sometimes divides itself. In Fig. 9 there are two sketches made by Graham.

The Crookes dark space (or cathode dark space, or Hittorf dark space as it is called in Germany) extends from the cathode to the boundary of the bright luminous cloud which is the negative glow. The boundary is generally so well-defined and distinct that an observer finds it easy to judge when a sounding-wire just touches it, or the cross-hair of a telescope coincides with its image; "in the case of oxygen," Aston said, "the sharpness was simply amazing; even with so large a dark space as 3 cm., the sighter could be set (to the boundary) as accurately as to the cathode itself, *i.e.*, to about 0.01 mm." I reproduce some of Aston's photographs in Figure 8, although he says that for reasons of perspective the boundary of the negative glow appears more diffuse than it really is.[17] The electric field strength within the Crookes dark space is greater, often very much greater, than in any of the other divisions of the discharge; almost the whole of the voltage-rise from cathode to anode is comprised within it, and the remainder, although spread across all the brilliant parts of the glow, is inconsiderable unless the tube is made unusually long. The behavior of the dark space when the current through the tube is varied (by varying a resistance in series with the tube) is curious and instructive. If the current is small and the cathode large (a wide metal plate) the negative glow overarches a small portion of the cathode surface, lying above it like a canopy with the thin dark sheath beneath it. When the current is increased the canopy spreads out, keeping its distance from the metal surface unaltered, but increasing its area proportionally to the current; the thickness of the Crookes dark space and the current-*density* across it remain unchanged. If the experimenter continues to increase the current after the cathode is completely overhung by the glow, the dark space thickens steadily, and the current-density across it rises.

The changes in the voltage across the Crookes dark space which accompany these changes in area and thickness are very important. The voltage is measured with a sounding-wire, like the cathode-fall in the arc; but since the boundary of the dark space is so sharply marked, the experimenter can set the sounding-wire accurately to it instead of merely as close as possible to the cathode. So long as the

[17] Adjacent to the cathode a thin perfectly dark stratum can be distinguished (especially in the picture on the right). The P.D. across this thin black space is, as nearly as it can be guessed from the width of the space, of about the magnitude of the ionizing-potential of the gas. In fact Aston estimated it for helium (to which the pictures refer) as 30 volts, a good anticipation of the value 24.5 assigned years later to the ionizing potential. It seems therefore that the outer edge of the very dark space is at the level where the electrons coming from the cathode first acquire energy enough to ionize.

negative glow does not overarch the whole cathode, and the thickness and current-density of the dark space keep their fixed minimum values, the voltage across it remains constant likewise. This is the *normal cathode-fall* of the glow. It is an even more thoroughgoing constant than the thickness or the current-density of the dark space, for these vary with the pressure of the gas (the dark space shrinks both in depth and in sidewise extension, if the current is kept constant while the gas is made denser) while the normal cathode-fall is immune to changes in pressure. It depends both on the gas and on the material of the cathode; the recorded values extend from about 60 volts (alkali-metal cathodes) to about 400 volts. Attempts have been made to correlate it with the thermionic work-function of the cathode metal, and there is no doubt that high values of the one tend to go with high values of the other, and low with low. When the cathode is entirely overspread by the negative glow and the dark space begins to thicken, the voltage across it rises rapidly; the cathode-fall is said to become *anomalous*, and may ascend to thousands of volts.

Almost the whole of the voltage-rise from cathode to anode, as I have stated, is generally comprised in the cathode-fall; the remainder, although spread over all of the brilliant divisions of the discharge, is inconsiderable unless the tube is unusually long. The field strength in the Crookes dark space is also much greater than anywhere else in the glow. This is illustrated by the two curves in Fig. 9, representing the field strength in the discharges sketched above them. (For the region of the Crookes dark space, however, the curves are defective.) In the luminous clouds the electric force is feeble, and they in fact are not essential to the current-flow; if the anode is pushed inwards towards the cathode, it simply swallows them up in succession without interfering with the current; but the moment it invades the Crookes dark space, the discharge ceases unless the electromotive force in the circuit is hastily pushed up. The mechanism which keeps the glow alive lies concealed in the dark space.

One naturally tries to invent a mechanism resembling the one suggested for the arc: the cathode-fall serves to give energy to the electrons emerging from the cathode, so that they ionize molecules at the edge of the negative glow; and the ions fall against the cathode with energy enough to drive out new electrons. But the details are more difficult to explain. The cathode-fall gives much more energy to the electrons than they need to ionize any known molecule, so that apparently its high value is what the ions require to give them enough energy to extract electrons from the cathode. We can hardly argue that the electrons are thermionic electrons; the cathode does not grow hot enough; if it does, the cathode-fall suddenly collapses, and the glow is liable to turn into an arc. Expulsion of electrons from cold metals by ions striking them has been separately studied, but not sufficiently.

On the other hand, there is good evidence that the Crookes dark space, like those dark sheaths scooped out in the positive column

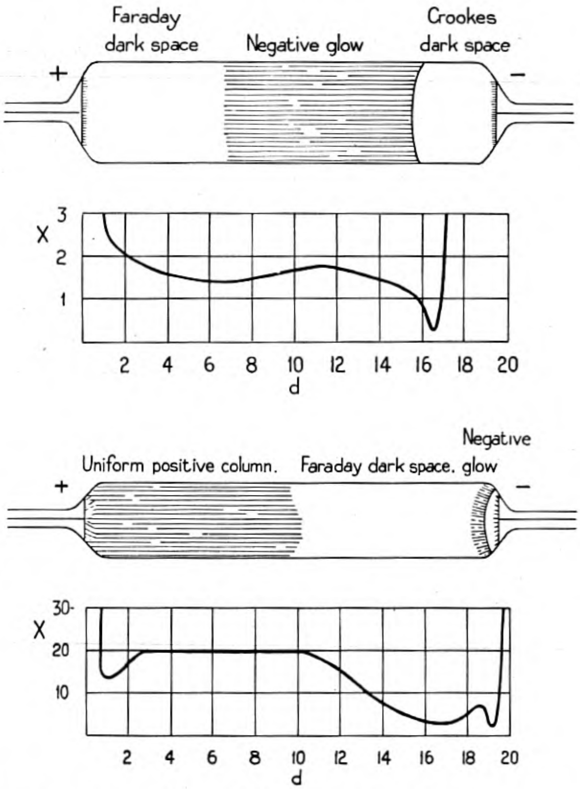


Fig. 9—Sketches of the glow in rarefied nitrogen at two pressures (the higher below) with curves showing the trend of field strength along the discharge. (Graham, *Wiedemanns Annalen*)

of the mercury arc by intruding a negatively-charged wire, is a region of predominantly positive space-charge, in which positive ions advance towards the cathode in a manner controlled by some such equation as (16) or (17). For example, Gunther-Schulze proposed (16) to describe the state of affairs in the Crookes dark space in the condition of normal cathode-fall; that is, he assumed that the ions fall unimpeded from the edge of the negative glow to the cathode surface. No doubt this assumption is too extreme, yet it leads to unexpectedly good agreements with experiment. Thus when the thickness of the Crookes dark space is altered (by altering the pressure of the gas) leaving the voltage across it constant, the current-density varies inversely as the square of the thickness, as it should by (16). And when Gunther-Schulze calculated the thickness of the dark space from (16), using the observed values of cathode-fall and current for six gases and two kinds of metal, and substituting the mass of the molecule of the gas for the coefficient m in that equation, the values he obtained agreed fairly well (within 40%) with the observed thicknesses. Long before, J. J. Thomson had proposed (17), and Aston tested it by a series of experiments on four gases, in the condition of strong anomalous cathode-fall. As k of that equation should be inversely proportional to the pressure p of the gas, the product id^3V^{-2} (V standing for the cathode-fall) should be constant at constant pressure, and the product $id^3V^{-2}p$ should be constant under all circumstances. These conclusions were fairly well confirmed for large current-densities.

Several attempts to test the theory by actually determining the potential-distribution in the Crookes dark space were made with sounding-wires and by other methods; but they have all been superseded, wherever possible, by the beautiful method founded on the discovery that certain spectrum lines are split into components when the molecule emitting them is floating in an intense electric field, and the separation of the components is proportional to the strength of the field. This was established by Stark who applied a strong controllable electric field to radiating atoms, and by LoSurdo who examined the lines emitted by molecules rushing through the strong field in the Crookes dark space, in the condition of anomalous cathode-fall. Now that the effect has been thoroughly studied it is legitimate to turn the experiments around and use the appearance of the split lines as an index of the field strength in the place where they are emitted. Brose in Germany and Foster at Yale did this. In the photographs (Fig. 10, 11) we see the components merged together at the top, which is at the edge of the negative glow, where the field is very small; thence they diverge to a maximum separation, and finally approach one another very slightly before reaching the bottom, which is at the cathode surface.[18] This shows that the net space-charge in the Crooke

[18] The displacements of certain components are not rigorously proportional to the field, and sometimes entirely new lines make their appearance at hitherto unoccupied places when a strong field is applied. Both of these anomalies can be detected in the pictures. For the original plate from which Fig. 11 was made I am indebted to Dr. Foster.

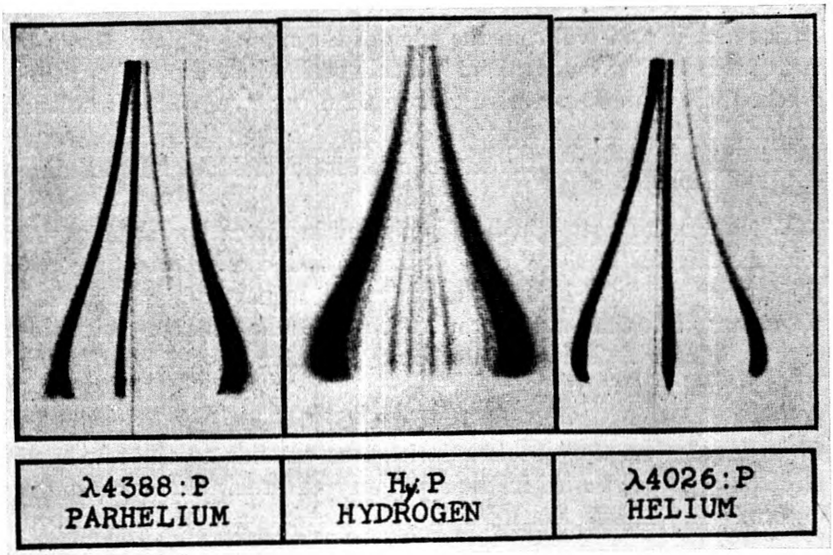


Fig. 10—Spectrum lines subdivided and spread out in the Crookes dark space by the strong and variable field. See footnote [18]. (J. S. Foster, *Physical Review*)

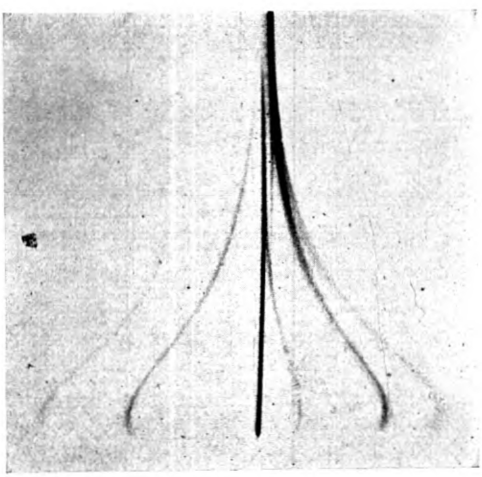


Fig. 11—A group of lines near λ4388 (parhelium spectrum) resolved and spread out in the Crookes dark space. See footnote [18].

dark space is positive from the edge of the negative glow almost but not quite to the cathode; there is a thin region just above the cathode where there is more negative charge than positive. This is splendid material for the theorist, and it is deplorable that the method cannot be applied except when the cathode-fall is anomalous and exceedingly large.

When a narrow straight hole is pierced in the cathode, the positive ions making for it shoot clear through, and can be manipulated in a chamber provided behind the cathode. In particular the ratios of their charges to their masses can be measured, and thence their masses can be inferred. This is Thomson's "positive-ray analysis," which Aston developed into the most generally available of all methods for analyzing elements into their isotopes. If the density of the gas is so far reduced that the Crookes dark space extends to the anode, the electrons can be studied in the same way and their charge-mass ratio determined. Hence the mass of the electron can be deduced, and its dependence upon the speed of the electron ascertained, yielding precious evidence in support of the special or restricted theory of relativity. These are among the simple phenomena which I mentioned at the beginning of this article, in which the properties of the ultimate atoms of electricity and matter are revealed.

The positive column, which is the brilliant, colorful and conspicuous part of the glow, resembles in some ways the positive column of the mercury arc. In it the potential-gradient decreases with increasing current, and the characteristic of the glow is negative (Fig. 6). Often the positive column subdivides itself into a regular procession of cloudlets or *striations*, all just alike and equally spaced (Fig. 7). The potential-difference between two consecutive striations has the same value all along the procession, and everyone feels instinctively that it ought to be the ionizing-potential or the resonance-potential of the gas; but this is evidently too simple an interpretation for the general case, although striations at potential-intervals of 4.9 volts have been realized in mercury vapor. Generally, if not always, the striations appear when the gas is contaminated with a small admixture of another. In this fact the key to the problem of their origin probably lies.

The *Glow in a dense gas* (as dense as the atmosphere, or more so) is visible only when the surface of either or both electrodes is curved, with a radius of curvature smaller than the minimum distance between the two. In these circumstances the field strength varies very greatly from one point to another of the interspace, at least before the space-charges begin to distort the field, and presumably afterwards as well;

it attains values just in front of the curved electrode (or electrodes, if both are curved) so great that if they prevailed over an equal interspace between flat electrodes they would instantly provoke an explosive spark. In some cases the glow in a dense gas resembles a very con-

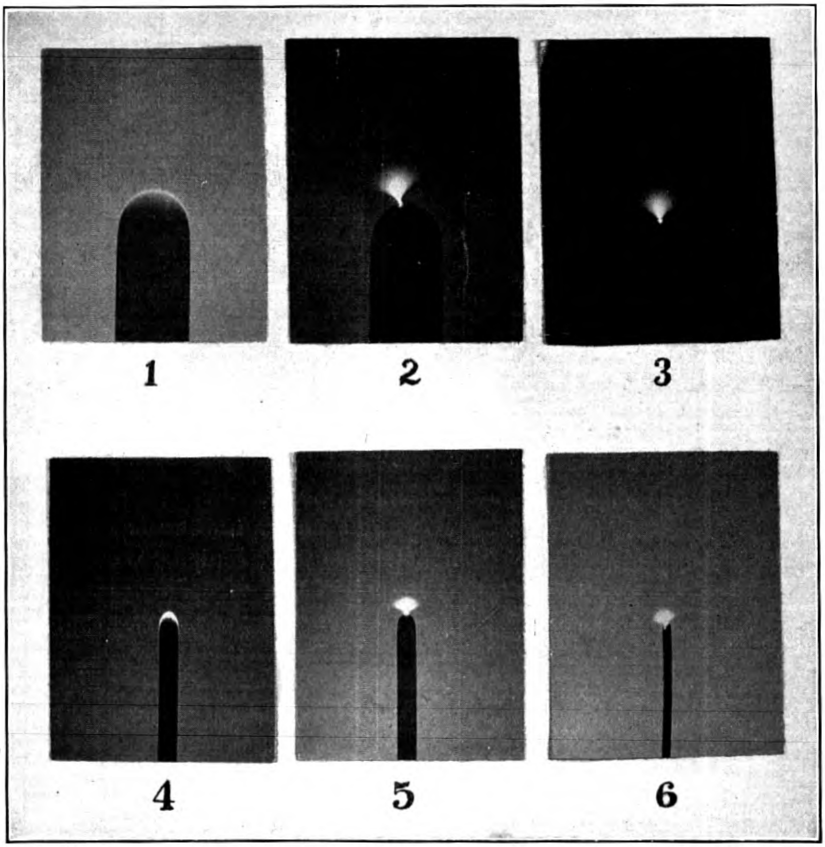


Fig. 12—The glow in air at atmospheric pressures, near a curved electrode (the other electrode is a plate beyond the top of the picture). In 1, 4 the curved electrode is the anode; in 2, 3, 5, 6 it is the cathode. (J. Zeleny, *Physical Review*)

tracted and reduced copy of portions of the glow in a rarefied gas. Thus in the photographs (Figs. 12, 13) of the luminosity surrounding a very curved cathode, it is possible to discern two dark spaces and two bright ones, the first dark space lying just outside the cathode, the last bright region fading off into the darkness which extends away towards the flat anode (far above and out of the picture). In the pictures of the glow surrounding a very curved anode, we see only a luminous sheath spread over the metal surface (Fig. 12).[19] Mathematically the simplest case (at least before the space-charge begins to affect the field) is realized by a slender cylindrical wire stretched along the axis of a

Fig. 13—Magnification of one of the pictures in Fig. 11. (The lowest bright spot is a reflection in the cathode surface)

much wider hollow cylinder, the wall of which may be imagined to recede to infinity in the limiting case. In this case the glow bears the euphonious name of *corona*, and has been intensively studied because it wastes the power transmitted over high-tension lines.

[19] I am indebted to Professor J. Zeleny for plates from which these figures were made.

Often there is a luminous cylindrical sheath encasing the wire, and from the boundary of the sheath outwards to the outer cylinder the gas is dark. It is customary to assume that the dark region, like the other dark spaces we have considered, is traversed by a procession of ions of one sign, positive or negative as the case may be, moving at a speed proportional to the field and controlled by their own space-charge according to the equation in cylindrical coordinates corresponding to (17); and the experiments support this assumption to a certain extent.

I must use my last paragraph to erase the impression—inevitably to be given by an account so short as this, in which the understood phenomena must be stressed and the mysterious ones passed over—that the flow of electricity through gases is to be set down in minds and books as a perfected science, organized, interpreted and finished. Quite the contrary! there are as many obscure and mysterious things in this field of physics as there are in any other which has been explored with as much diligence. Its remarkable feature is not that most or many of the phenomena in it have been perfectly explained; but rather, that for those few which have been explained, the explanations are very simple and elegant; they are based on a few fundamental assumptions about atoms and electrons which are not difficult to adopt, for they are not merely plausible but actually *demonstrable*. Perhaps as time goes on all the phenomena will be explained from these same assumptions. There will be experimenters who modify the apparatus and the circumstances of past experiments so that all of the avoidable complications are avoided and the phenomena are simplified into lucid illustrations of the fundamental principles; and there will be theorists, who take the complicated phenomena as they are delivered over to us, and extend the power of mathematical analysis until it overcomes them. They may find it necessary to make other and further assumptions, beyond those we have introduced; at present it is commonly felt that ours may be sufficient. Whether posterity will agree with us in this, must be left for posterity to decide.

Carrier Telephony on High Voltage Power Lines

By W. V. WOLFE

Introduction

THE use of power from hydro-electric generating stations and central steam plants has increased until single companies serve a territory of many thousands of square miles and the problem of coordinating the distributing centers with the generating stations has steadily increased in complexity.

One of the essentials of this coordination is obviously an adequate system of communication and until the recent advent of high frequency telephony, this service was secured over privately owned telephone lines and over lines of public service telephone companies.

The advent of the power line carrier telephone system now offers a highly reliable and satisfactory means of communication in connection with the operation of power systems. This equipment has been designed to employ the power conductors as the transmission medium and to provide service as reliable as the power lines themselves with a low initial cost, a small maintenance charge, increased safety for the operating personnel and transmission comparable in quality and freedom from noise with that obtained on high grade commercial toll circuits.

Preliminary Problems

In proceeding with the development of the Western Electric Power Line Carrier Telephone System three major problems were encountered. It was first necessary to learn from field tests and close contact with power companies the characteristics of power lines and associated apparatus at high frequencies and the operating requirements for such a telephone system; second, it was necessary to develop a safe and efficient method for coupling the carrier apparatus to the power conductors and third, to select and develop circuits and equipment suited to this service.

The superiority of the full-metallic over the ground return high frequency circuit was easily established by comparative measurements of attenuation, noise and interference, and therefore the experimental work was largely confined to the former circuit.

High Frequency Attenuation of Power Lines

Since the measurement of the attenuation of a circuit ordinarily requires that the circuit be terminated in its surge impedance [1] to avoid reflection effects, the first step in determining the attenuation of the power line was to measure its surge impedance. After considering several methods for measuring this impedance, a substitution

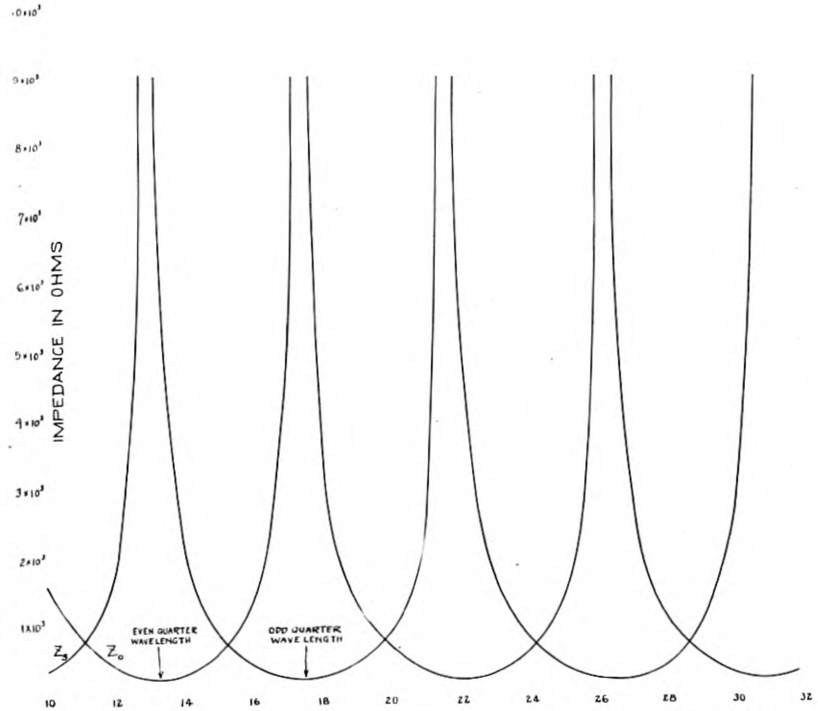


Fig. 1—Open Circuit (Z_o) and Short Circuit (Z_s) Impedance as Measured at Carrier Frequencies on a 110,000 Volt Power Line 12 Miles Long

method was adopted because of its simplicity and the rapidity with which measurements could be made. This method depends upon the fact that the apparent or measured impedance of a uniform line terminated in its surge impedance is equal to that surge impedance and it consists in terminating the line in a known resistance and determining the value of current supplied to the line by an oscillator

[1] Surge or characteristic impedance may be defined as the measured impedance of a uniform line of infinite length or in the case of a finite line it may be expressed mathematically as $Z = \sqrt{Z_{\substack{\text{open}\\ \text{circuited}}} \times Z_{\substack{\text{short}\\ \text{circuited}}}}$

and then substituting for the line a non-inductive resistance until the same value of current is drawn from the oscillator. In employing this method for determining the surge impedance it was assumed that the oscillator output was constant, and that the phase angle of the surge impedance was small.

A study of the curves on Fig. 1 shows that the apparent impedance of the line will change with the impedance in which the line is terminated in different ways, depending upon the frequency used. (1) If

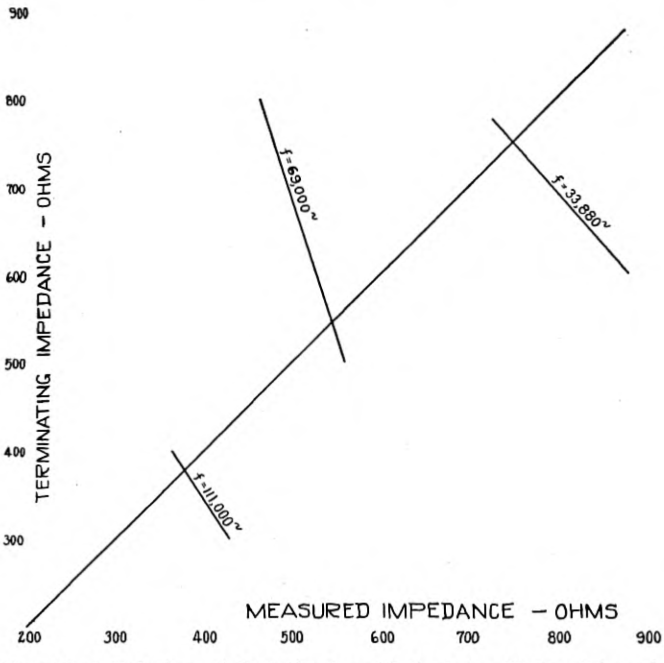


Fig. 2—Graphical Solution of Substitution Method for Determining the Surge Impedance of a Power Line

a frequency mid-way between the quarter wave lengths [2] is used, the open circuit and short-circuit impedances are the same. (2) If a frequency corresponding to an even quarter wave length is used, an increase in the terminating impedance will produce an increase in the apparent impedance of the line. (3) If a frequency corresponding to an odd quarter wave length is used, an increase in the terminating impedance will produce a decrease in the apparent imped-

[2] Whenever the length of the line becomes equal to, or some multiple of, one quarter of the length of the electric wave of the corresponding frequency, it is referred to as a quarter wave length frequency, or, for short, a quarter wave length.

ance of the line. If the apparent impedance of the line is plotted against the terminating impedance, in (1) the curve will be horizontal; in (2) the curve will have a positive slope approaching 45° and in (3) the curve will have a negative slope of approximately 45°. Each of these curves will intersect a 45° line drawn through the origin at a point where the terminal impedance is equal to the surge impedance of the line. This intersection can be determined with the

Fig. 3—Frequency vs. Attenuation and Frequency vs. Surge Impedance as Measured on the Tallulah Falls-Gainesville 110,000 Volt Power Line

greatest ease and accuracy when the curve crosses the 45° line at right angles or under condition (3), that is, when the determination is made at a frequency corresponding to an odd quarter wave length. To determine the surge impedance at a given frequency all that was necessary was to terminate the line at the distant end in an impedance which it was anticipated would be just below the surge impedance and measure by the substitution method the apparent impedance of the line, and then to terminate the line at the distant end in an impedance which would just exceed the surge impedance and determine the corresponding apparent impedance. The intersection of a straight line through these points with the 45° line determined the correct terminating impedance. In Fig. 2 is shown a determination of the characteristic impedance of the Tallulah Falls-Gainesville line of the Georgia Railway and Power Company at three different frequencies.

The attenuation of the line was then measured by terminating it in its characteristic impedance and measuring the current in to the line and current out of the line.[3] The results of the attenuation measurements made on the Tallulah Falls-Gainesville line are shown on Fig. 3. The irregularities in the attenuation shown by the lower

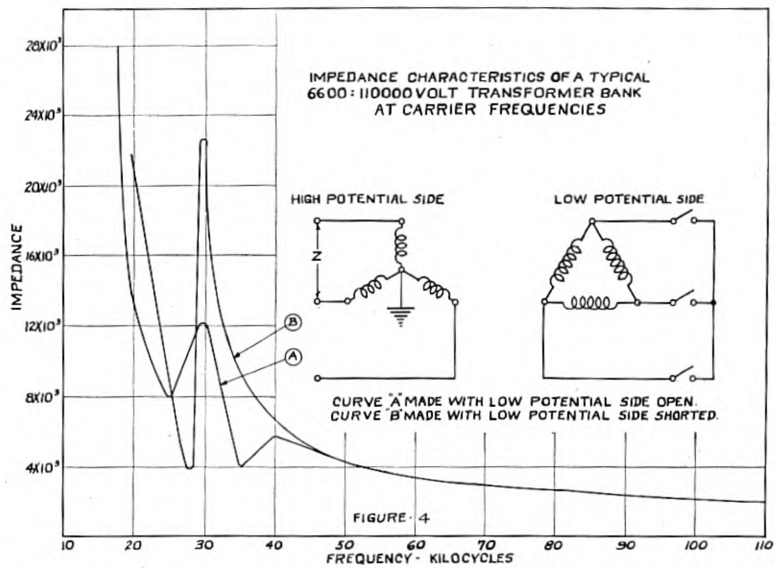


Fig. 4—Impedance Characteristics at Carrier Frequencies of a Typical 6600:110000 Volt Transformer Bank

curve are probably caused by the error in assuming that the phase angle of the surge impedance was small and that the surge impedance was a straight line function of frequency. From these and other data it was evident that for frequencies as high as 150 K.C. the attenuation is not excessive.

High Frequency Characteristics of Power Transformers

In order to determine the effect of power transformers on the use of the power line as a transmission medium for high frequency currents,

[3] Attenuation expressed in transmission units is equal to 20 $\log_{10} \frac{I_1}{I_2}$ where I_1 is the current into the network and I_2 is the current received from the network and measured in a circuit whose impedance corresponds to the characteristic impedance of the network.

the impedance of typical transformer banks was measured. In Fig. 4 is shown the impedance versus frequency characteristic of a three phase, 110,000–6600 V., 12,000 K.V.A. transformer bank connected "star" on the high side with the neutral grounded and "delta" on the low side. As shown by the diagram, these measurements were made between phases on the high side with the low side open circuited and short circuited. The coincidence of these curves for frequencies above 50 K.C. indicates that at these frequencies the dominant characteristic is the distributed capacity of the high winding and the impedance is probably unaffected by changes on the low potential side of the transformer. Below 50 K.C., however, the impedance changes rapidly both with frequency and with the low potential termination.

A study of Figs. 3 and 4 and other data shows that the desirable frequency range in which to operate a power line carrier telephone circuit is that from 50 K.C. to 150 K.C. In this range the attenuation is not excessive, it is very little affected by the associated power apparatus, and it is independent of the conditions on the low potential power circuits. The curve shown in Fig. 3 indicates that, contrary to the common belief, the attenuation in this range is a relatively smooth function of frequency. This conclusion is supported by the fact that in the various installations of power line carrier telephone equipment which have been made since the attenuation measurements on Fig. 3 were obtained, no power lines have been encountered where the attenuation was a critical function of frequency. Another important argument for the selection of this frequency range lies in the fact that it is well above the range employed for multiplex telephony on commercial telephone systems and therefore precludes any interference with such systems.

Coupling Between Carrier Equipment and Power Line

Probably the most difficult problem to solve was that of providing a satisfactory method for connecting the carrier equipment to the power line. The use of power transformers has not been found practicable for if frequencies low enough to be efficiently transformed were employed, the attenuation of the circuit would be a function of the conditions in the distributing network and a change in the number or arrangement of transformers would result in an appreciable change in the attenuation. Such a method of coupling to the power line would also have the objection that communication would not be possible when the power transformers were disconnected from the line.

Since it did not seem practicable to develop a carrier frequency transformer suitable for connecting between phases of a high voltage power line it was decided to couple to the power line by means of capacity. Two general types of condensers are possible, first, a concentrated capacity condenser and second, a distributed capacity condenser. A concentrated capacity condenser suitable for direct connection to a high voltage power line was not available, but its development has been successfully undertaken by the Ohio Brass Co.

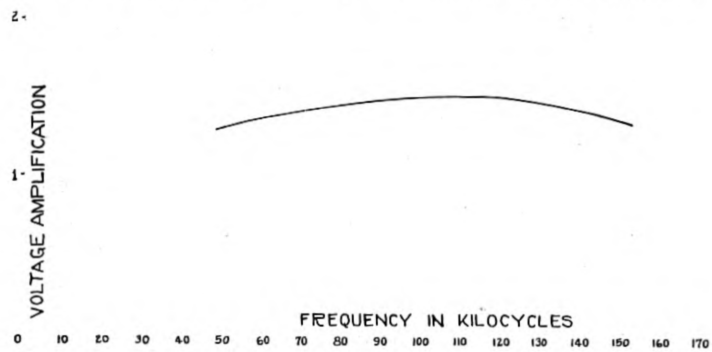


Fig. 5—Voltage Amplification Characteristic of High Frequency Transformer

The distributed capacity was obtained by suspending a wire parallel to the power conductor and employing this wire as one plate of the condenser and the conductor as the other plate. Both of these methods of connecting to the power line have been developed and are described later.

Design of the Carrier Equipment

Although the "carrier suppressed" system has many advantages over the "carrier transmitted" system, the difficulty of securing filters suitable for suppressing the unwanted products of the modulation prevented the use of the carrier suppressed system.

Several general characteristics of the electrical and mechanical design of this carrier equipment are worthy of note. The various stages of vacuum tubes in both the transmitting and receiving circuits are coupled by transformers. These transformers are closed iron core coils using the standard core employed for audio-frequency transformers. Fig. 5 shows the characteristic of one of these transformers, and it is evident from this figure that the variation in amplification from 50 K.C. to 150 K.C. is only a fraction of a transmission unit.

Although the frequencies employed by this equipment are fairly high, it was practicable to mount all of the apparatus on standard steel relay rack plates. In order to minimize the maintenance on this equipment no "C" batteries have been employed, the grid potentials

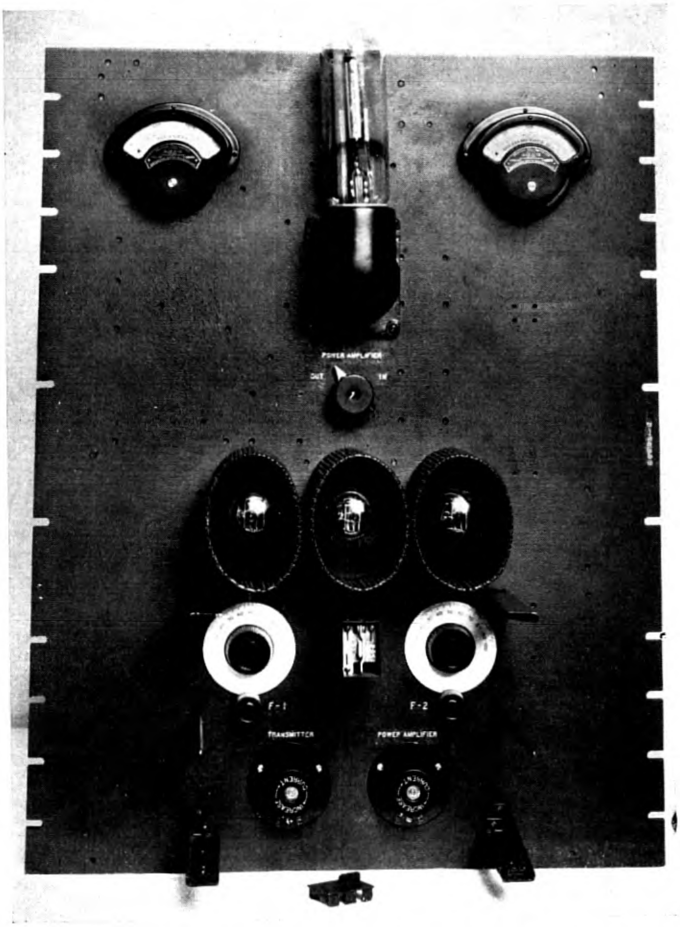


Fig. 6—Front View of Transmitter Panel with Cover Removed from Tuning Condensers

being obtained from filament drop, "B" battery drop and a combination of these two.

The transmitting unit shown in Figs. 6 and 7 is divided into two parts, the transmitting circuit proper and the power amplifier. The first is a circuit comprising a 101-D tube functioning as a Hartley oscillator with inductive feed-back, a 223-A tube operating as a speech amplifier or modulator and a 223-A tube operating as a high frequency amplifier. The plate or constant current system of modulation is employed but differs somewhat from the usual practice in

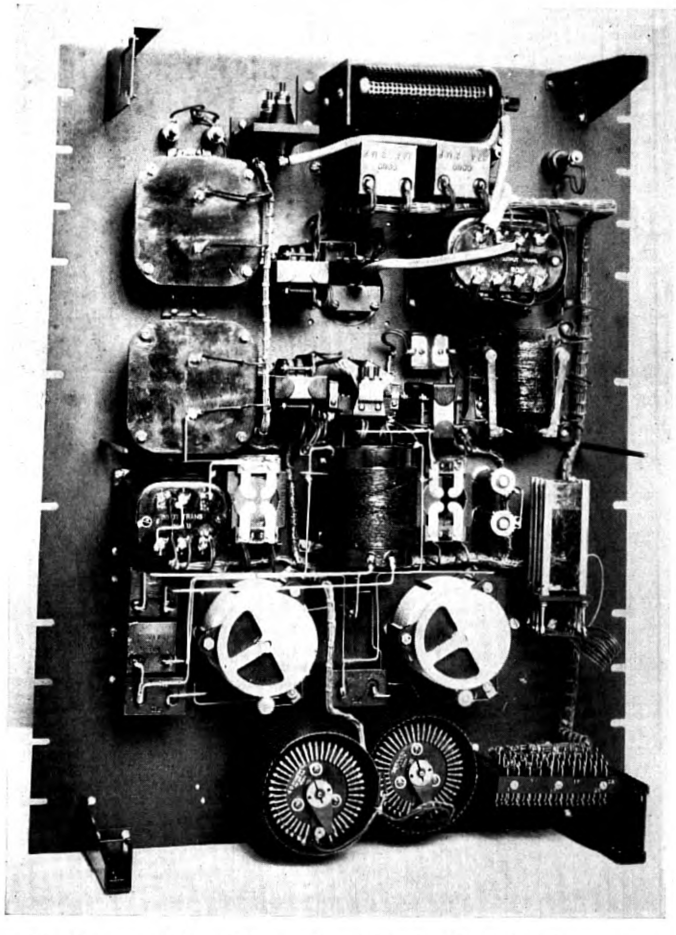


Fig. 7—Rear View of Transmitter Panel with Cover Removed

that the output of the high frequency amplifier is modulated rather than the output of the oscillator itself. This scheme was found to deliver more modulated power than the usual arrangement since it is not limited to the same extent by the overloading of the high frequency amplifier. This circuit has a power output of one watt, which has proved to be ample for normal operation of the carrier system.

To provide for operation of the system when the attenuation on the power line has been materially increased by line fault conditions a power amplifier is provided. This amplifier employs a 50 watt tube (211-A) and is placed in the circuit by a simple switching operation. When this amplifier is operated, the output of the transmitting circuit is impressed upon the grid of the 50 watt tube and amplified to approximately fifty times its normal power output.

In the present type of carrier system duplex or two way operation is secured by the use of two different carrier frequencies, one for transmission in each direction. As will be pointed out later in the

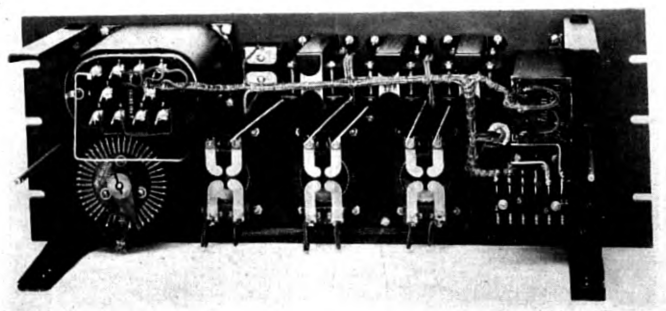


Fig. 8— Rear View of Receiver Panel

section on signaling the lower frequency is always assigned to the calling station. The transmitting circuit must therefore operate at two different frequencies. This change is accomplished by the automatic operation of the relay shown in Fig. 6. The operation of this relay changes the capacity in the oscillating circuit, thereby changing its frequency. The values of the two frequencies at which the transmitting circuit operates are determined by the variable condensers F1 and F2, Fig. 6, and certain fixed condensers which are connected in parallel with the variable condensers.

The receiving unit shown in Fig. 8 is extremely simple. It is not tuned and the only control is the filament rheostat. It consists of three 101-D vacuum tubes operating respectively as a carrier frequency amplifier, a negative grid potential detector and an audio frequency amplifier.

Two way operation is secured by operating the transmitting and receiving circuits at different frequencies and separating them by means of filters. In the single channel systems this separation is secured by a high pass filter and a low pass filter although in the multiple channel system band pass filters will be employed. Fig. 9 shows attenuation versus frequency characteristics of the high and low pass filter combination. A study of these curves shows that the transmission loss or attenuation in the high pass filter to frequencies transmitted by the low pass filter is never less than 90 T.U., which corre-

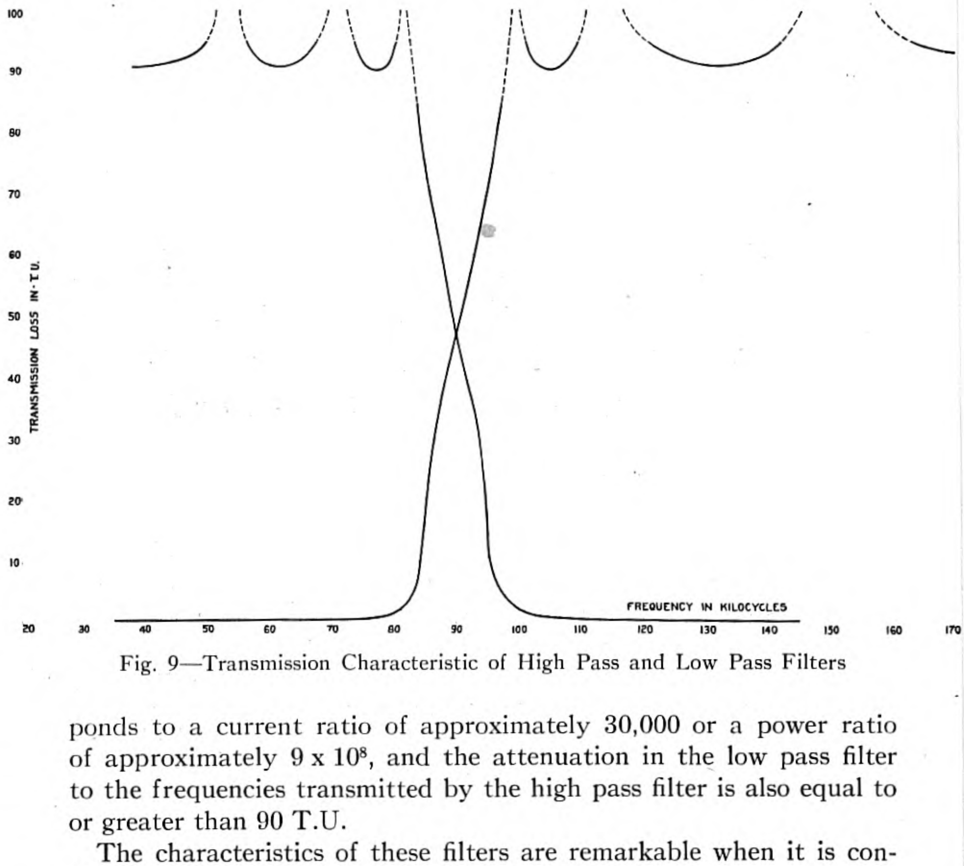


Fig. 9—Transmission Characteristic of High Pass and Low Pass Filters

ponds to a current ratio of approximately 30,000 or a power ratio of approximately 9×10^8, and the attenuation in the low pass filter to the frequencies transmitted by the high pass filter is also equal to or greater than 90 T.U.

The characteristics of these filters are remarkable when it is considered that the frequency range in which they operate is higher than that employed for multiplex carrier telephone systems, the attenuation secured is higher than that ordinarily required for such systems, and a power of 50 watts has to be transmitted through them thereby introducing special problems in the design of the coils and condensers. Figs. 10 and 11 are front and back views of one of these filters.

One of the unusual features in the use of these filters is the fact that the position of the filters in the circuit is changed from time to time by the operation of the relay shown on Fig. 11, that is to say, when the transmitting circuit is operating at a frequency lower than 80 K.C. the low pass filter is connected to it and when the transmitting circuit is operating at a frequency higher than 100 K.C. the high pass filter must be connected to it.

Signaling System

Signaling or ringing is accomplished at the transmitting end by changing the frequency of the oscillator from a frequency below 80 K.C. to a frequency above 100 K.C. without changing the filters. This is

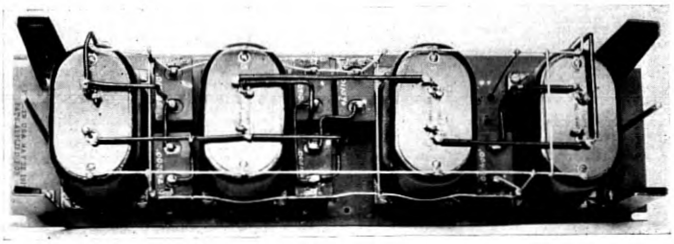


Fig. 10—Front View of Low Pass Filter with Cover Removed

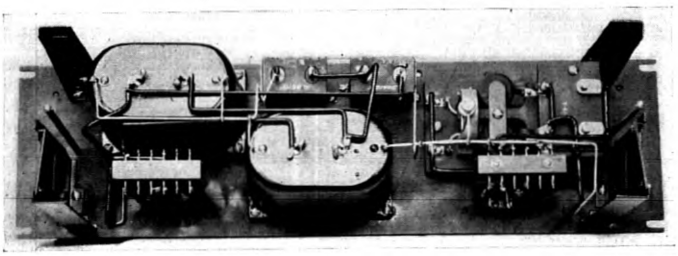


Fig. 11—Rear View of Low Pass Filter with Cover Removed

accomplished by operating and releasing the relay in the oscillator circuit. Since the filter connected to the transmitting circuit will pass only one of these frequencies, pulses of the carrier frequency are sent out on the line. At the receiving end these pulses are amplified and rectified and the change in the space current of the detector operates a marginal relay. The number and arrangement of these pulses is controlled by a spring-operated selector key of the type commonly employed for telephone dispatching on railroad lines. At the receiving end these pulses operate a train dispatching selector relay (see Fig. 12) which responds to 17 impulses. This selector relay will respond to only two arrangements of these 17 pulses. The first arrangement is 17 consecutive pulses in which case these pulses must follow one another at the correct speed and must be of the correct duration. This makes it possible to ring all stations at the same time as may be desirable in issuing general orders. The selector relay will also respond to 17 pulses broken up into three groups in which case the correct number of pulses must occur in each group and the total of the three groups must be 17. This makes it possible to

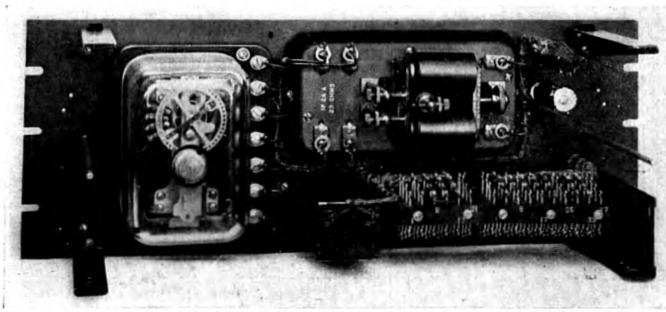


Fig. 12—Rear View of Signaling and Low Frequency Panel Showing the Signaling Apparatus

select one station from a group of more than 50 stations without disturbing the others. In addition to these desirable characteristics a single selector relay will provide selective ringing on four low frequency extensions from the carrier terminal.

The carrier equipment may be operated with complete control and talking facilities from either a telephone located at the carrier terminal or a telephone some distance from the carrier terminal but connected to it by a physical telephone circuit. In any event the control is automatic, the transmitting circuit operating only when the receiver is off the switchhook, while the receiving circuit operates continuously.

Designating the carrier frequency which is below 80 K.C. as F_1 and the carrier frequency which is above 100 K.C. as F_2, the operation of a carrier system comprising three carrier terminals designated as A, B and C with a remote control station designated as A_1 located at the load dispatcher's office and separated from the carrier terminal by several miles of physical telephone circuit is as follows. Each of these stations may communicate with any of the other stations. Communication between A, B and C is carried on over carrier circuits; communication between A and A_1 is carried on over the physical circuit while communication between A_1, B and C is carried on over circuits which are composed of a carrier circuit and a physical circuit operating in tandem. When in the normal or non-operated conditions, each of these carrier terminals is set up to receive a signal on frequency F_1, but when the receiver is removed from the switch-hook at any station to initiate a call, the carrier terminal corre-

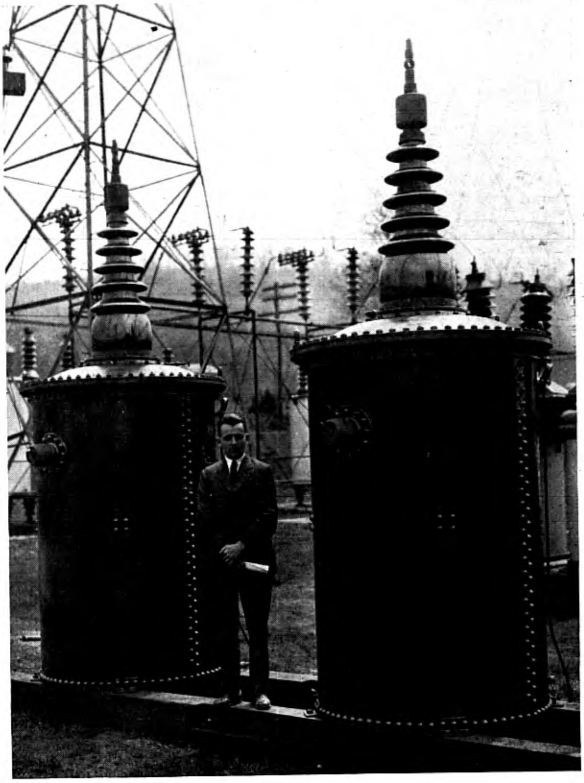


Fig. 13—110 K.V. Coupling Condensers Used for Coupling Carrier Circuit to a 110 K.V. Power Line

sponding to that telephone is automatically set up to transmit on frequency F_1 and receive on frequency F_2. When the ringing key is operated, pulses of frequency F_1 are sent out and received at all of the other carrier terminals. At the called station these pulses operate a selector relay and ring the bell, and when the operator removes his receiver from the switch-hook to answer the call, his carrier terminal is automatically set up to transmit on frequency F_2 and receive on frequency F_1. This switching of the transmitting and receiving circuits from one frequency to another is necessary where more than two stations are operated on the same system and it is desirable for every station to be able to call every other station without routing the call through a central point.

If station A_1 is connected with station A by means of two or more pairs of telephone wires which are not exposed to high voltage power

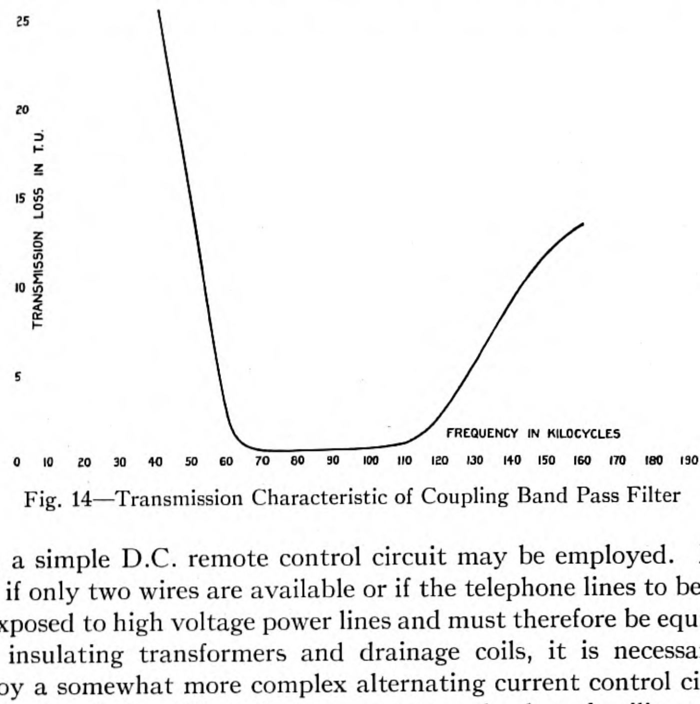


Fig. 14—Transmission Characteristic of Coupling Band Pass Filter

lines, a simple D.C. remote control circuit may be employed. However, if only two wires are available or if the telephone lines to be used are exposed to high voltage power lines and must therefore be equipped with insulating transformers and drainage coils, it is necessary to employ a somewhat more complex alternating current control circuit. In this circuit the 135 cycle interrupters and relays familiar to the telephone plant are employed.

The voice frequency circuits used in connection with this carrier equipment are the standard two wire and four wire circuits used in commercial telephone practices.

Coupling by Condensers and by Distributed Capacity

Fig. 13 shows two of the 120 K.V. coupling condensers developed by the Ohio Brass Co. Each of these condensers has a capacity of .003 μf although similar condensers having a capacity of .007 μf are also available. These condensers are approximately 5 ft. in diameter and 12 ft. high over the bushing and weigh about 8,000 pounds. The condenser element is made up of a large number of small condensers in parallel, the assembly being immersed in transformer oil.

At present these condensers are employed as the series capacity element of a single section, confluent type, Campbell band pass filter as shown by Fig. 22, the general attenuation characteristic being shown by Fig. 14. This filter is intended to transmit efficiently the carrier frequencies, and to exclude power frequency currents.

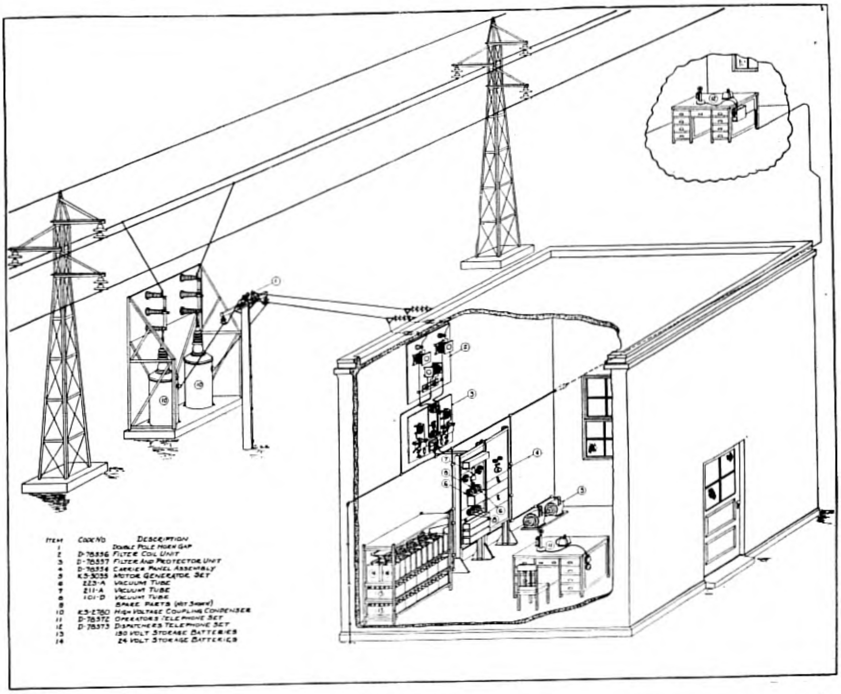


Fig. 15—Typical Layout of Power Line Carrier Telephone System, Using High Voltage Condensers for Coupling to Power Line

In Fig. 15 is shown a typical layout of a condenser coupled power line carrier telephone system.

In employing the distributed capacity type of condenser for coupling to the power line, two coupling wires (sometimes incorrectly called antennae) are suspended parallel to the power conductors for a distance of approximately 1,000 ft. Fig. 16 shows the last tower supporting the coupling wires in an installation at Anniston, Alabama. This is a twin circuit 110 K.V. power line and in order to secure coupling to both lines, the coupling wires are suspended midway between the top and bottom phases. The box shown on the tower in Fig. 16 is the coupling wire tuning unit shown in Fig. 17. The coupling wires are terminated in this tuning unit. In Fig. 18 is

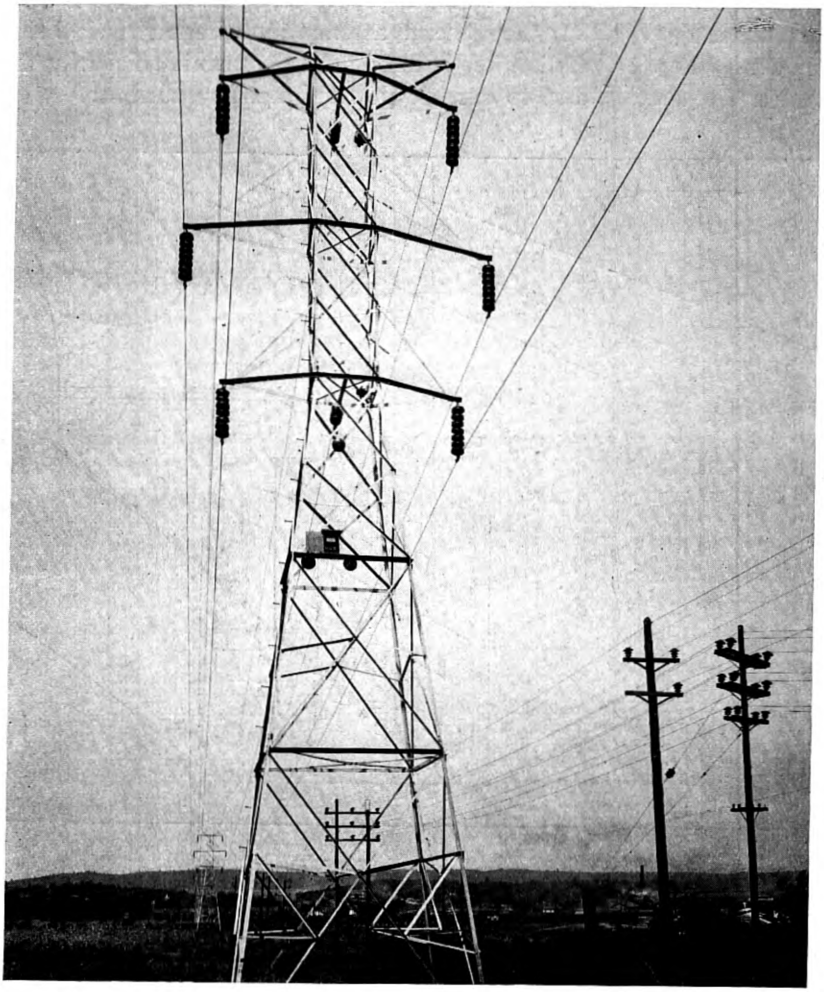


Fig. 16—Distant End of Typical Coupling Wire Installation Showing Coupling Wire Tuning Unit

shown the schematic diagram of the wire coupling circuit and Fig. 19 illustrates the character of the resonant peaks secured by this circuit. The series inductances L_1 and the terminating inductance L_2 are variable and by adjusting them the points of resonance may be shifted to correct for variations in the coupling wire inductance and capacity for different installations. Fig. 20 illustrates a typical carrier terminal installation employing the wire coupling method.

The only point in favor of the wire coupling as compared with the condenser coupling is the fact that for power lines of voltages higher

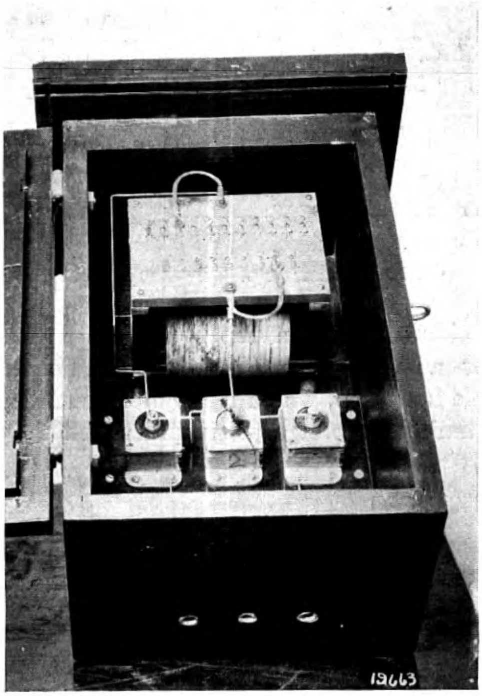


Fig. 17—Coupling Wire Tuning Unit

than 33 K.V. it is somewhat cheaper. On the other hand condenser coupling is much more efficient, thereby increasing the range and reliability of the system. It also permits high quality transmission, the transmission through it is not affected by small variations in frequency, and the component parts are of constant value determined at the time of manufacture and require no adjustment at the time of installation. In addition to these advantages the inspection and maintenance of the condenser is easier than for the coupling wires.

Protective Measures

In considering the problem of safety to the operating personnel and the equipment from the power line voltage, the normal insulation supplied by the high voltage condenser where it is employed or by the air separation where the coupling wires are employed, is disregarded, since this insulation may fail, thereby applying the power line voltage to the line terminals of the coupling circuit shown in

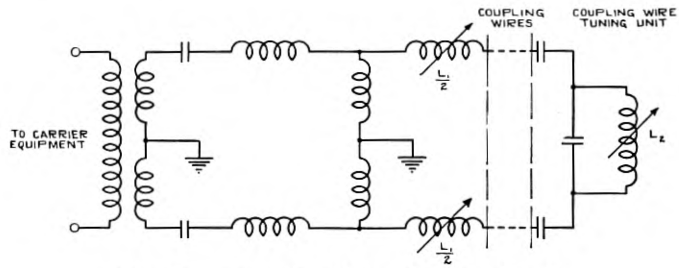


Fig. 18—Schematic of Wire Coupling Circuit

Fig. 21. The circuit shown in this figure is the same both for condenser and for wire coupling installations. The first element of protection is the horn gap, which is mounted outside of the building and serves to limit the voltage to ground which the drop wire fuse,

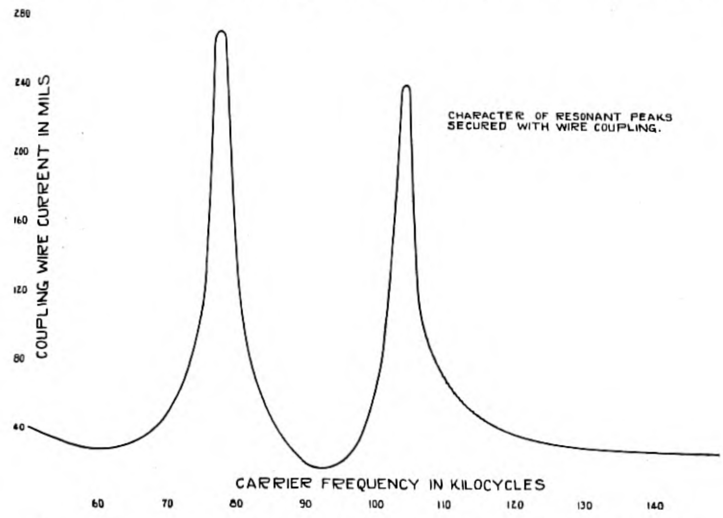


Fig. 19—Character of Resonant Peaks secured with Wire Coupling

constituting the second element of protection, will have to break. This fuse consists of an element inside of a porcelain tube the ends of which are closed by lead caps. This fuse is about 5 inches long and ½ inch in diameter and is supported by the wire itself. When it fails, the arc established within the porclain tube causes the tube to break and permits the wires to fall apart. In power line carrier telephone practice this fuse is so installed that a clear drop of at least 20 ft. is obtained. The third element of protection is the shunt coil with the mid-point grounded. In many respects this element is the

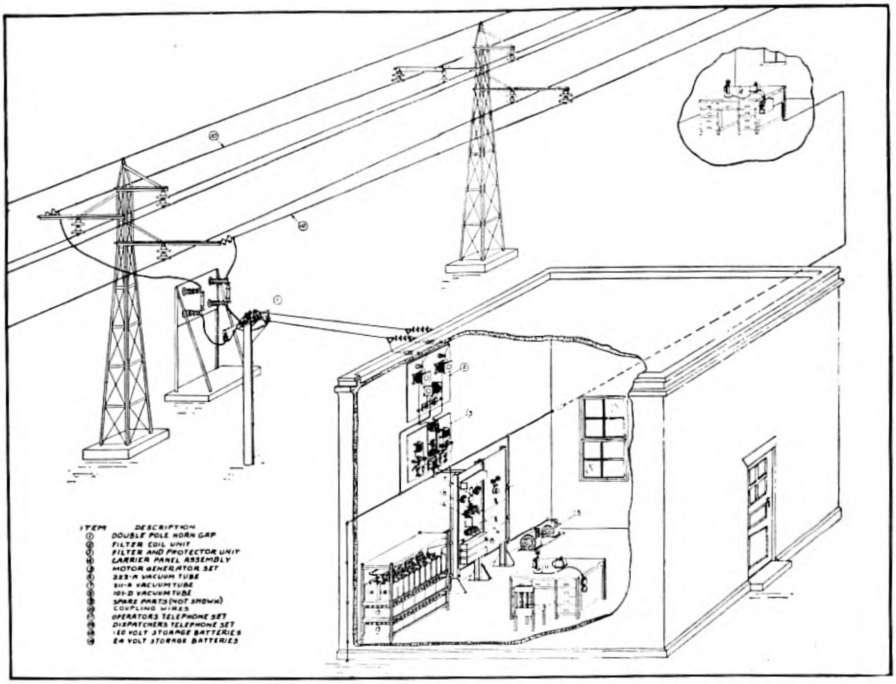


Fig. 20—Typical Layout of Power Line Carrier Telephone System Using Wire Coupling

most important one, since it provides a low impedance path to ground for power frequencies, thereby draining off the 60 cycle potentials which are collected by either the coupling wires or the condensers in normal operation.

As will be noted from Fig. 27 the line series inductances and this shunt inductance coil comprise a unit (the upper panel) which is known as the filter coil unit. The coils on this unit are insulated for 20,000 volts on the line terminals and are constructed of edgewise wound copper ribbon large enough to carry heavy momentary currents without damage. The fourth element of protection is a fused switch and surge arrester such as is commonly employed for the protection of private telephone lines exposed to power lines. This device consists of fuses in series with the line and forming the blades of a switch. These fuses have been found satisfactory for the interruption of voltages as high as 25,000. Following this fused switch is a 1,500 volt breakdown static spark gap to ground and a 500 volt breakdown vacuum gap

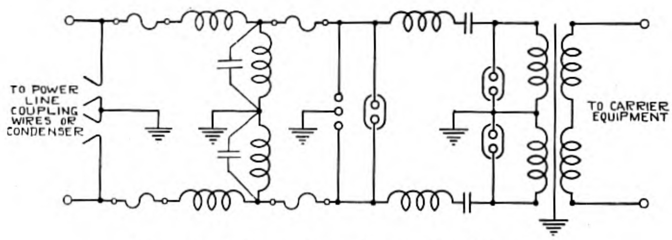


Fig. 21—Schematic of Protection Circuits

across the line. Following these there are two series capacity elements which are high voltage mica condensers. These condensers have a capacity of .007 μf. and a breakdown voltage in excess of 7,500. Finally, there is provided a repeating coil with the mid-point of the line side

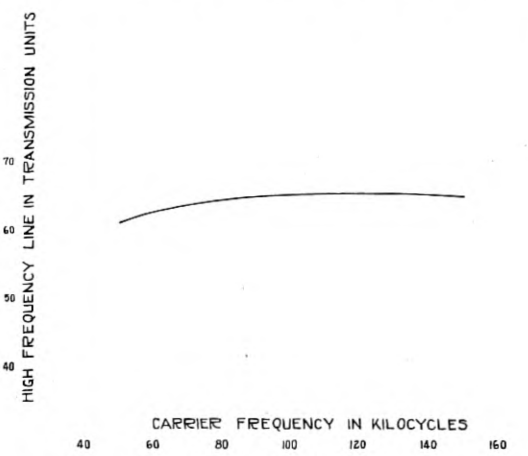


Fig. 22—Change in the Attenuation of the High Frequency Line Necessary to Maintain a Constance Voice Frequency Level with Variation in the Frequency of the Carrier

winding grounded and protected by 500 volt vacuum gaps to ground. This repeating coil is also provided with a grounded shield between the windings and has a breakdown voltage from the winding to the shield of 1,000 volts. The operation of this protective circuit has been demonstrated several times in the field by connecting one phase of a 110 K.V. power line directly to one of the line terminals of the protective circuit. In every case the circuit has operated satisfactorily. In no case has any of the standard apparatus been damaged nor has there been any evidence that the elements of protection beyond the third, that is, the shunt coil with the mid-point grounded, have been called upon to function.

Transmission Level Characteristics

Fig. 22 shows the attenuation (expressed in transmission units) of the high frequency line versus the carrier frequency of K.C. It will be noted that over the range from 50 K.C. to 150 K.C. the variation

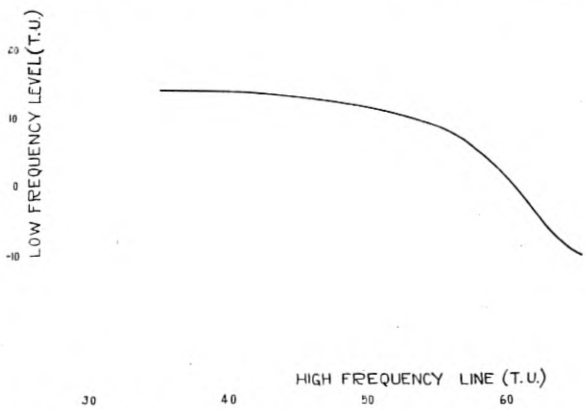


Fig. 23—Variation of Overall Gain with the Attenuation of the High Frequency Line

in attenuation is less than 5 T.U. This curve was made with a constant audio frequency input of 3.35 mils and an output of 3.35 mils from the carrier circuits, the audio frequency being 1,000 cycles. The variation of audio frequency level with the attenuation of the high frequency line is shown in Fig. 23. The observations given in Fig. 24 were made on an artificial transmission line in which the line constants, and therefore the attenuation, could be readily changed without changing the carrier frequency. The shape of this curve is a function of the receiving circuit since the audio input, carrier frequency and the modulated output of the transmitting circuit are maintained constant. It shows that for audio frequency levels lying between -10 and $+10$ T.U. the equivalent is approximately a straight line function of the attenuation of the high frequency line, and that therefore the receiving circuit is not overloaded.

Fig. 24 shows the audio frequency load characteristic. This curve is principally a function of the load characteristic of the modulator and it shows that for inputs greater than 1 mil, the modulator is overloaded. In practice the overloading of the modulator is prevented by increasing the average low frequency line equivalent to an attenuation of 10 T.U. by means of a resistance artificial line. This

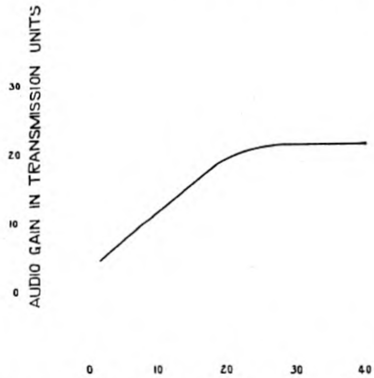


Fig. 24—Transmitting Circuit Load Characteristic

arrangement is desirable in order that the balancing of the low frequency hybrid coil may not be complicated when operating over very short physical circuits.

The curve in Fig. 25 is a single frequency quality characteristic and shows that where the method employed for connecting to the

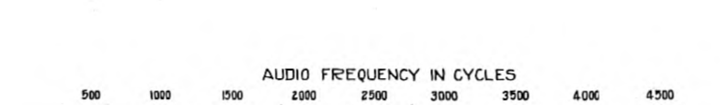


Fig. 25—Single Frequency Quality Characteristic

power line will permit, remarkably true voice transmission may be secured. The variation in the equivalent over the range from 100 cycles to 5,000 cycles is only 5½ T.U., while the variation from 300 cycles to 5,000 cycles is only 2 T.U. Reference to Fig. 19 will indicate, however, that less satisfactory quality characteristics are obtained when the wire coupling method is employed, because of the sharpness of resonance of the coupling circuit.

Alabama Power Company Installation

Figs. 26 and 27 are photographs of the installation of power line carrier telephone equipment at the Anniston substation of the Alabama

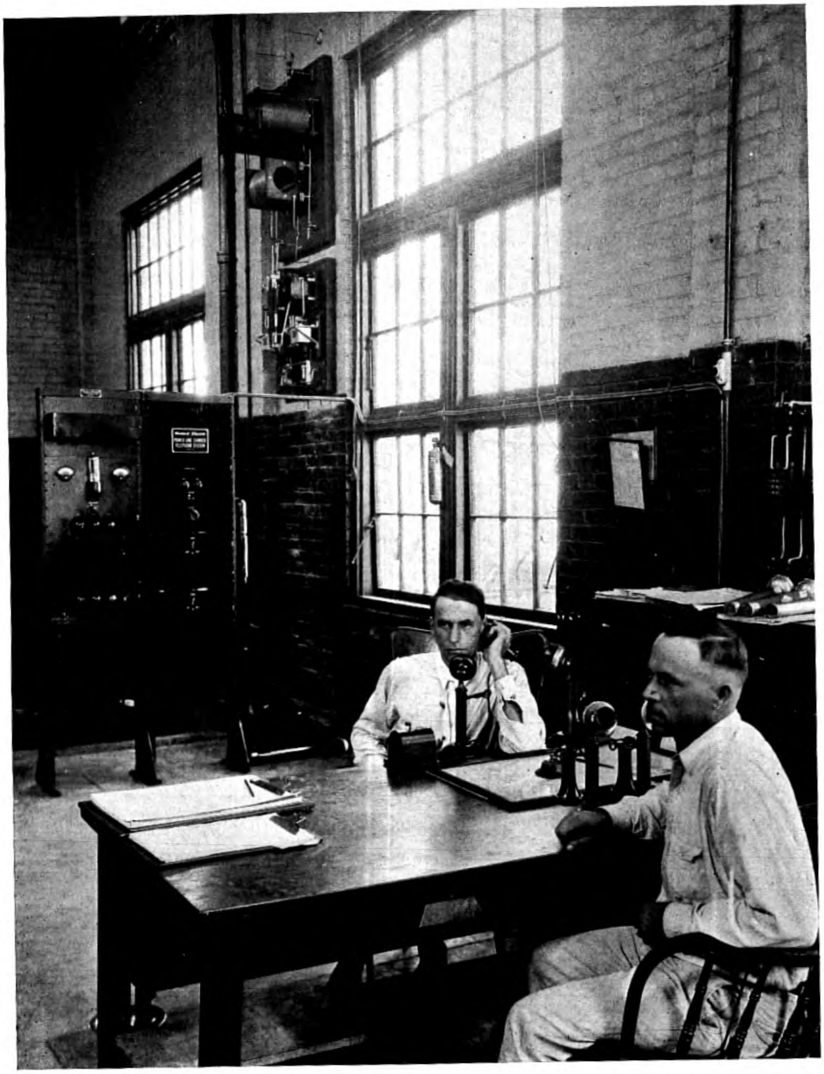


Fig. 26—Typical Power Line Carrier Telephone Installation

Power Company. Fig. 26 illustrates the simple character of the assembled units and freedom from controls. The right hand bay is devoted to power control apparatus with space reserved for the 135 cycle remote control equipment when it is employed. The left

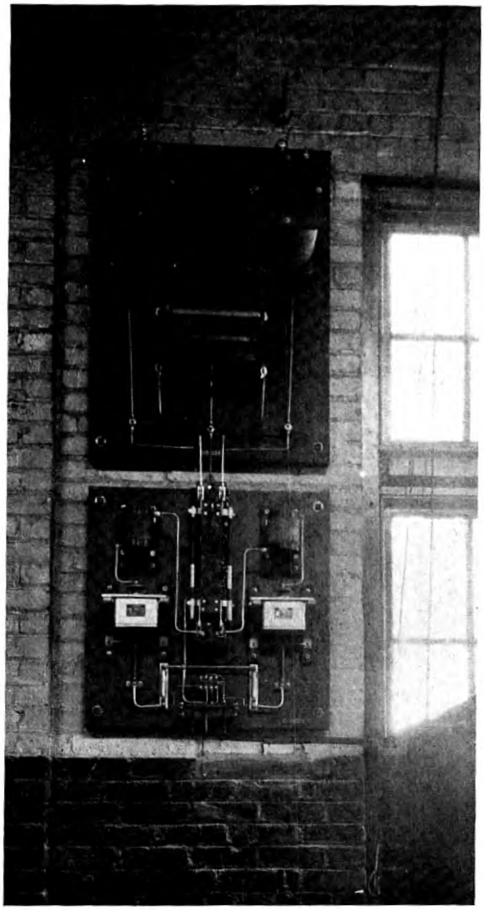


Fig. 27—Typical Installation of Coupling Panels

hand bay includes the transmitting and receiving circuits, the high and low pass carrier frequency filters and the voice frequency and D.C. control circuits. Beginning at the top of this bay, the first panel, which is blank on front, carries the system terminal. block to which all wiring except the power supply is connected. The second panel is the high pass filter; the third panel is blank. The fourth panel is the transmitting equipment, both low power and high power. The fifth panel is the receiving circuit; the sixth panel contains the voice frequency and signaling equipment. The seventh panel contains D.C. control equipment, and the bottom panel is the low pass filter. On the wall to the right of the carrier panel assembly are shown the filter coil unit and the filter and protector unit. These units are more clearly shown in Fig. 27 and diagrammatically in Fig. 21. Returning to Fig. 26, the desk stand which the operator is using is that associated with the carrier equipment, while the key mounted on the table immediately to the left of the desk stand is the selector key employed for ringing. Fig. 16 shows the coupling wire installation at this station.

The power line carrier telephone equipment which has been briefly described in the foregoing article is in successful operation today on several power systems in this country. Its reliability, simplicity of operation and maintenance have been well established.

The large number of variables which are involved in line failure conditions make it impossible to predict what effect these emergency conditions may have on the operation of the carrier equipment. The fact remains, however, that under many simulated and actual trouble conditions successful operation of the carrier equipment has been obtained.

With the growing need of power companies for communication facilities, it is probably only a question of a very short time before multiple channel carrier systems will be in operation on the large power systems of this country.

Abstracts of Bell System Technical Papers Not Appearing in the Bell System Technical Journal

Photomechanical Wave Analyzer Applied to Inharmonic Analysis.[1] C. F. SACIA. This type of Fourier Analysis deals with wave-forms which are not strictly periodic, since they are of finite duration and of varying cyclic forms. Hence in a finite frequency range they have an infinite number of infinitesimal components (shown by the Fourier Integral) as contrasted with the finite number of finite components at regular intervals (shown by the Fourier Series).

This analyzer utilizes the continual repetition of the aperiodic wave, deriving therefrom a periodic wave, the infinitesimal components neutralizing except for frequencies which are integral multiples of the frequency of repetition; here the components build up to finite magnitudes. The simple relation between these components is seen from the corresponding Fourier Integral and Series identities for the unrepeated and repeated waves respectively. By increasing the period of repetition a new set of components can be similarly derived.

The wave form is represented as a black profile on a transparent strip whose ends are joined to form an endless belt. Driven at constant speed past a transverse illuminated slit, it generates light fluctuations which are converted into electrical fluctuations by means of a selenium cell. A tuned circuit, amplifier, rectifier and microammeter are used to select and measure the components, while the frequencies are determined by the speed of the strip, the frequency of tuning, and the time scale of the original wave form.

"*Demagnetization and Hysteresis Loops.*"[2] L. W. MCKEEHAN and P. P. CIOFFI. The fact that permalloy shows its maximum initial permeability in the absence of external magnetic fields is used to check the exact compensation of the earth's magnetic field or other stray fields by measurement of the initial permeability of a strip or wire of permalloy placed parallel to the field component to be compensated. Increased accuracy is obtained by the use of somewhat greater fields than those which approximately give the initial permeability. The effect of demagnetization by an alternating current field is studied with samples of the same sort, the apparent permeability varying as the external field at the time of magnetization is varied. The dissymmetry in hysteresis loops where the upper and lower limits

[1] J. O. S., R. S. I., Vol. 9, pp. 487–494, 1924.

[2] J. O. S., R. S. I., Vol. 9, pp. 479–485, 1924.

are unsymmetrical with respect to the zero of magnetic field is illustrated and the detection of such dissymmetry is discussed.

A Classified List of Published Bibliographies in Physics, 1910–1922.[3] KARL K. DARROW. This work, undertaken at the request of the National Research Council, represents an attempt to cope with the problem of providing a convenient and adequate bibliography of physics, not by actually writing a complete classified bibliography (which would fill a huge volume and require the prolonged labor of several men), but by listing the very numerous partial bibliographies under a detailed subject-classification. Many of the accounts of research published in scientific journals contain short histories of the previous work in the subjects which they treat, many others contain lists of references, and there are also a number of critical or uncritical reviews of particular fields with thorough documentations. The *Classified List of Published Bibliographies* refers to all of these which appeared in any of the familiar physical journals between 1910 and 1922 inclusively, and a number of books as well; it is believed that almost every article upon a physical subject, which has ever been cited or reviewed in another article, can be traced through the *List*. The system of classification, in which the field of physics is divided into seventy-five classes with numerous subdivisions, is much the most detailed and elaborate which has been made out for the science of physics in a score of years. An adequate system of classification is of great value in any science, for researches which are clasified under it are not only made easy to trace, but their various aspects and their mutual relations can be emphasized. Because of the rapid growth and evolution of physics, the earlier systems have mostly become inadequate; but it is hoped to make and keep this system effective by constant attention and revision, and to extend the use of it.

Transmitting Equipment for Radio Telephone Broadcasting.[4] EDWARD L. NELSON. The general transmission considerations applying to any system for the high quality transmission of speech or music are outlined briefly, and the specific requirements to be met by the various apparatus units in a radio broadcasting equipment are discussed in some detail. The standard Western Electric 500-watt broadcasting equipment, which has found application in some fifty of the larger stations in this country and abroad, is described. Its performance capabilities are illustrated and it is indicated that a standard of performance has been attained which renders possible reproductions not substantially different from the original.

[3] Bulletin of the National Research Council, No. 47.

[4] Proc. of The Inst. of Radio Engineers, Vol. XII, page 553, 1924.

"The Vapor Pressures of Rochelle Salt, the Hydrates of Sodium and Potassium Tartrates and Their Saturated Solutions."[5] H. H. LOWRY and S. O. MORGAN. The vapor pressures were determined by a static method at several temperatures between 15° and 40°. Temperatures were controlled to ±0.1° and the pressures read to ±0.1 mm. The measurements on the saturated solution of Rochelle salt show that the solid phase in such a solution is unstable above 40°, in agreement with other investigators.

Minimal Length Arc Characteristics.[6] H. E. IVES. This paper is a study of the electrical discharges which occur between opening contacts. It is found that the discharge occurring when currents below a certain value are broken are atmospheric sparks corresponding to a definite breakdown voltage, which in the case of air is about 300 volts. Above a critical value of current, which is different for every material, the discharge is an arc, in which the voltage corresponding to the discharge varies with current. Spectograms taken in the two regions show only the air spark spectrum for all materials below the critical current and the arc spectra of the materials above the critical current. The characteristic equations of the arcs caused by the opening contacts are derived and are used to obtain expressions for the current vs. time relations at the opening contact.

The Dependence of the Loudness of a Complex Sound Upon the Energy in the Various Frequency Regions of the Sound.[7] H. FLETCHER and J. C. STEINBERG. Two complex sounds were studied, one with a continuous energy frequency spectrum corresponding to connected speech, the other a test tone having discrete frequency components. By means of filters the energy was removed from all frequencies either above or below a certain frequency, and the resulting decrease in loudness was measured by attenuating the original sound without distortion until equal in loudness to the filtered sound. Taking the average results for six observers, this decrease was found to depend on the absolute values of the loudness. For a loudness of 22 units above threshold, each frequency region contributes to loudness in proportion to the energy in that region weighted according to the threshold energy for that frequency. For a loudness above 30 units, however, this is no longer true, because of the non-linear character of the response of the ear. By assuming each frequency region contributes in proportion to a fractional power of the weighted energy of that region, values of the total loudness in agreement with ob-

[5] Jour. Am. Chem. Soc., Vol. 45, pp. 2192–2196, 1924.

[6] Journal of the Franklin Institute, Vol. 198, pp. 437–474, 1924.

[7] Physical Review, Vol. 24, page 306, 1924.

served values are obtained if proper values are taken for the fractional power, decreasing to one third as the loudness increases to 100 units.

Correlation Between Crack Development in Glass While Conducting Electricity and the Chemical Composition of the Glass.[8] EARLE E. SCHUMACHER. A study was made of the susceptibility to crack development shown by five different kinds of glass when they were subjected to the action of an electric current. The results indicated that the tendency to crack increased with increasing alkali content of the glass and with increasing electrical conductivity.

Report of the Chairman of the Telegraphy and Telephony Committee of the American Institute of Electrical Engineers.[9] O. B. BLACKWELL. This report gives a brief summary of the advances which have been made or which have come into prominence in the communication art during the year. Papers which have been presented before the Institute and which, in general, have recorded such advances are reviewed.

Selective Circuits and Static Interference.[10] J. R. CARSON. This paper is an application of a general mathematical theory to the question as to the possibilities and limitations of selective circuits when employed to reduce "Static" interference. In the case of static interference and random disturbances in general the random and unpredictable character of the disturbances makes it necessary to treat the problem statistically and express the results in mean values. In spite of the meagre information available regarding the character and frequency distribution of static, this treatment of the problem yields general deductions of practical significance. The conclusion is reached that for given signal requirements there is an irreducible residue of static interference which cannot be eliminated. This limit is closely approached when a filter of only two or three sections is employed as the selective circuit, and only a negligible further gain is made possible by the most elaborate circuit arrangements. A formula is also given for calculating the relative figures of merit of selective circuits with respect to random interference.

The Guided and Radiated Energy in Wire Transmission.[11] J. R. CARSON. This is a mathematical analysis of wave propagation along guiding wires from the fundamental equations of electromagnetic theory. It is shown that the engineering theory of wire transmission is incomplete, and that, in addition to the transmitted wave of en-

[8] Jour. Am. Chem. Soc., Vol. XLVI, No. 8, August, 1924.

[9] Journal of the American Inst. of Elec. Engineers, Vol. 43, page 1083, 1924.

[10] Trans. A. I. E. E., 1924.

[11] Jour. A. I. E. E., Oct., 1924.

gineering theory, an infinite series of complementary waves exist. It is through these waves that the phenomena of radiation are directly accounted for. Except for the phenomena of radiation, however, the complementary waves are of theoretical rather than practical interest in present-day transmission practice, and except in extreme cases they may be ignored in practice without appreciable error.

Sound Magnification and Its Application to the Requirements of the Deafened.[12] HARVEY FLETCHER. A general description of the generation and propagation of sound waves was given and experiments performed to illustrate the principles involved. The general requirements for aiding persons having various amounts of deafness were outlined. The relation between the loudness of speech received by the ear in a room of average acoustic characteristics and the distance the speaker is away from the ear was illustrated by a chart. Also, a chart showing the characteristic frequency regions and loudness levels of the fundamental speech sounds, and one showing the interpretation of speech at various loudness levels by persons having various degrees of hearing, were exhibited. By means of these three charts it was shown how one could predict the amount of intelligibility which would be obtained by a person having a definitely measured amount of hearing. In particular it was pointed out that such sounds as *th*, *f*, and *v* will be the first sounds to be lost as the hearing decreases. These sounds are the easiest ones to detect by lip reading so that hearing aids and lip reading go hand in hand in aiding one who is hard of hearing to obtain the proper interpretation.

Abstract of a Telephone Transmission Reference System.[13] L. J. SIVIAN. The subject is dealt with in four parts: A—The function of a transmission reference system; B—Requirements to be met by the reference system; C—Work done on the construction and calibration of a preliminary model of the new reference system; D—Proposed future development of the new reference system in its final form to be adopted as the standard for the Bell System.

A brief discussion of the methods and apparatus entering into the general problem of rating telephone transmission is given. It is

[12] Lecture given before the Annual Conference of the American Federation of Organizations for the Hard of Hearing, Washington, D. C., Thursday, June 5, and published in Volta Review, September, 1924.

A large number of the audience who listened to this lecture were hard of hearing. A rough measurement of the amount of hearing of each of those present was made and groups arranged according to the degree of hearing. The amplification was then adjusted to each group to suit their particular needs. The results seemed to be most gratifying, as nearly everybody said that it was the first time they ever heard a public lecture of this sort without difficulty since they had become hard of hearing.

[13] Electrical Communications, Vol. III, pp. 114–126, 1924.

concluded that a physical reference system is essential, and that a mere specification of its physical operating characteristics is insufficient. The inadequacy of the reference systems now in use is pointed out.

The conditions to be aimed at in the new reference system are: I—The performance of the system and of its component parts must be specifiable in terms of quantities admitting of definite physical measurement; II—The performance of the reference system, under specified operating and atmospheric conditions, must remain constant with time; III—The reference system must be free from non-linear distortion over the range of acoustic and electric amplitudes which it must handle; IV—The frequency response over the range of speech frequencies must be as nearly uniform as possible.

Of the above, conditions I and II are regarded as the most important. It is also proposed to build auxiliary reference systems which will meet conditions I and II while falling short of III and IV. These are needed for purposes of ready comparisons with the commercial circuits commonly in use.

Contributors to this Issue

F. L. Rhodes, S.B., Massachusetts Institute of Technology, 1892; American Bell Telephone Company; Outside Plant Engineer, American Telephone and Telegraph Company, 1909–19; Outside Plant Development Engineer, 1919—. Mr. Rhodes has had an active part in the development and standardization of materials, apparatus and practices employed in the underground and overhead wire plant of the Bell System. He has written many articles, among which may be mentioned those on "The Telephone" in the Encyclopedia Americana and Nelson's Encyclopedia.

George Crisson, M.E., Stevens Institute of Technology, 1906; instructor in Electrical Engineering, 1906–10. American Telephone and Telegraph Company, Engineering Department, outside plant division, 1910–14; transmission and protection division, 1914–19; Development and Research Department, transmission development division, 1919—.

W. H. Harden, B.E.E., University of Michigan, 1912; Engineering Department, American Telephone and Telegraph Company, 1912–1919; Department of Operation and Engineering, 1919—. Mr. Harden has been engaged in the development of transmission maintenance testing methods and in the preparation of routines and practices required for applying these methods in the telephone plant.

K. S. Johnson, A.B., Harvard University, 1907; Graduate School of Applied Arts and Sciences, 1907–09; Engineering Department of the American Telephone and Telegraph Company, 1909–13; Engineering Department, Western Electric Co., Inc., 1913–24; Bell Telephone Laboratories, Inc., 1925—. Mr. Johnson's work has related especially to the theoretical aspects of telephone and telegraph transmission.

Timothy E. Shea, S.M., Massachusetts Institute of Technology, 1919; instructor in Electrical Engineering and Physics, 1918–20; Manufacturing Department, Western Electric Company, 1920–21; Engineering Department, 1921–24; Apparatus Development Department, Bell Telephone Laboratories, 1925—. Mr. Shea has been principally engaged in the development of electric wave filters and allied apparatus.

Karl K. Darrow, S.B., University of Chicago, 1911; University of Paris, 1911–12; University of Berlin, 1912; Ph.D., in physics and mathematics, University of Chicago, 1917; Engineering Department, Western Electric Company, 1917–24, Bell Telephone Laboratories, Inc., 1925—. Mr. Darrow has been engaged largely in preparing studies and analyses of published research in various fields of physics.

W. V. Wolfe, B.S., Carnegie Institute of Technology, 1918; Signal Corps, 1918–19; General Electric Company, 1919; Standard Underground Cable Company, 1920; Engineering Department, Western Electric Company, 1920–24; Bell Telephone Laboratories, Inc., 1925—. Mr. Wolfe has been engaged in the development of various types of carrier systems.