The Bell System Technical Journal

October, 1930

Chemistry in the Telephone Industry[1]

By ROBERT R. WILLIAMS

An account is given of the activities of the Chemical Department of the Bell Telephone Laboratories. In the Laboratories chemists act chiefly as advisers and critics. They concern themselves with such problems as the theory of chemical structure as related to dielectric properties and simultaneously attack the task of making an improved substitute for gutta-percha which renders possible a transatlantic telephone. They are interested in the colloidal structure of cotton and silk and the influence of moisture and electrolytes on their insulating properties. The dispersion hardening and fatigue resistance of lead and its alloys, the fabrication of platinum-alloy vacuum-tube filaments, and of new magnetic materials such as permalloy and perminvar have required their attention. The corrosion of cable systems is due largely to stray currents but is strongly influenced by chemical factors. Other underground corrosion, and particularly the slower and more insidious corrosion and tarnishing of indoor telephone apparatus, have justified a broad program of investigation of the corrosive factors involved in ocean, earth, and atmosphere. Related to these studies are those of protective finishes, whether metallic coatings or organic paints and lacquers. The permanence of a great variety of materials must often be predicted as best it may without the test of service life. This interest in permanence is reflected in a program of experiments in preservation of telephone poles.

THE public mind associates the chemist with glass retorts and evil smells, with war gases or the glare of furnaces against the sky. To the technical leader in industry, however, the wide distribution of chemists outside predominantly chemical enterprises has become a familiar fact. A casual reference to the thirty subject classification headings in *Chemical Abstracts* will serve to illustrate how widely industrial chemists have become disseminated and how large a volume of work they are producing.

When one reflects that all engineering is essentially applied physics, and that it has become subdivided into a score of specialized fields, it seems very natural that chemistry also should have found varied applications as the sum total of chemical knowledge has increased. Perhaps the day is not far distant when the term "physical chemist" will represent to the lay mind as well-defined and distinct a calling as that of civil engineer.

It is therefore only a part of a general movement in industry that has placed the chemist in a position of some importance in the telephone business. His relative importance in the communication field is small, as the industry must permanently remain essentially electrical rather

[1] *Ind. & Engg. Chem.*, April, 1930.

than chemical. His usefulness depends primarily, not upon the number or size of the operations which are entrusted to his exclusive care, but upon the distinctive mode of thought which he contributes to a critical consideration of the methods and processes in use. As in medicine, so in telephony the chemist is an aid to progress, not a prime mover.

If one endeavors to define the distinctive mode of thought of chemists, he is at once led to point out the fact that the chemist by his training intuitively tries to account for most phenomena by a consideration of the composition of the materials involved. From his first days in the laboratory he is taught what "chemically pure" means and learns that even minor impurities may often have most important consequences, good or bad. While it is obviously true that many things happen without a chemical cause, it is equally, though less obviously, true that variations of composition or chemical changes in composition are frequently associated with the happenings. No one is so well qualified as the chemist to ferret out such obscure correlated and often important facts.

The distinctive nature of the training of chemists tends to adapt them to the role of critics. The fundamentals of the old-school physics reached a point some thirty or forty years ago when it was felt that the whole field had been fairly thoroughly combed over and all essential principles were known. These principles became embodied in formulas and conventionalized modes of attack upon the problems of engineering which have often been accepted at more than their true value by the average product of the engineering school. Chemistry, on the other hand, has never reached so high a development. Even in first-year chemistry one encounters facts not explained satisfactorily by any known theory, and the reading of even a score of pages of an elementary chemistry which has passed its first printing will bring one upon statements which require modification in the light of more recently acquired knowledge. By comparison with applied classical physics, chemistry is a youthful science and its devotees are inclined to a juvenile disrespect for tradition.

The enormous consumption in a telephone plant of such materials as lead and copper, paper and textiles, rubber and asphaltic compounds immediately implies the necessity of the chemist for the performance of his most conventional function, that of analysis. While the Bell Telephone Laboratories does not undertake the systematic inspection analyses of the large variety of products purchased, it does undertake a great volume of analysis of such products as a referee. Such work often reveals defects of analytical methods or defective statement of specification requirements that necessitate large numbers of comparative analyses to form a basis for proper amendment.

Another large portion of the work of our analytical laboratory has to do with materials used in the prosecution of research problems by physicists, engineers, and chemists in other groups. For example, in the course of development of such magnetic materials as permalloy and perminvar, methods for the analysis of unusual alloys have had to be devised. The utilization of microchemical and electrometric methods has often offered a way out of difficulties. For instance, it became necessary in connection with a corrosion problem to measure accurately the amounts of volatile acids in certain woods, and for this purpose a differential electrometric titration method was developed. Spectroscopic analysis finds its use, not only in the examination of minute specimens of material such as the deposits on vacuum tube filaments, but also in the estimation of minute impurities in grosser products, as, for example, the presence of zinc in solder. A not inconsiderable volume of research in analytical chemistry has grown out of these problems.

The most important function of the chemist in the telephone laboratory is not a conventional one. It consists in a scrutiny of the apparatus, equipment, materials, and processes of the industry to determine where and how chemical reactions or variations in composition are affecting functional operation. Sometimes the problems so encountered lead to extended researches pursued over a period of years to answer specific questions or to accumulate a reservoir of general information to be drawn upon as needs arise. Sometimes questions are brought to us by workers in other fields, apparatus designers perhaps, questions such as can be answered offhand, or at most require a few days or weeks of work. A considerable part of the more interesting work is published.

In order to make clear how chemistry applies to telephone problems, a number of examples have been chosen from various parts of the general field. Many of the tasks enumerated below are shared by the chemists with technical staffs of other departments. They will be discussed from the chemical viewpoint with only sufficient reference to the general engineering considerations to make the problems intelligible.

Chemical Constitution and Electrical Characteristics of Pure Substances

The most fundamental piece of work in dielectrics which we are undertaking is a study of chemical constitution of pure substances in relation to their electrical characteristics. This work is going forward under the direction of H. H. Lowry, and is expected to serve as a valuable supplement to similar work being conducted elsewhere, principally in university laboratories. While it is well known to all chemists that aqueous solutions are in general relatively good conductors, and that solutions of substances in fat solvents are relatively good insulators, a much closer analysis of the latter class is necessary for telephone purposes. Among the electrical characteristics of importance are dielectric constant, insulation resistance, a.c. conductivity, and dielectric strength. Each of these characteristics varies over quite wide ranges, depending upon certain conditions, of which temperature and frequency of alternation of the electric current may be mentioned as most important.

In this work an endeavor is being made to determine, for example, how the symmetry of the molecule affects the dielectric constant, and to distinguish in dielectrics the contribution of energy loss made severally by the electron, the atom, and the molecule. For such a purpose it is obviously necessary to deal with highly purified substances and to begin with those of simple chemical structure. Many of these, such as hexane, benzene, ethyl ether, and alcohol, are not expected to have the slightest importance as practical insulators, but serve to show how particular atomic groupings affect dielectric behavior. For this purpose we have included the methyl halides in our study, making measurements both in the pure state and in dilute solution, of dielectric constant and a.c. conductivity at temperatures ranging from 100° C. to the boiling point of the liquid. The frequency of the current has also been varied from 1 to 100 kilocycles.

Rubber and Gutta-Percha

Parallel with this study of the theory of dielectric behavior, a number of more immediately practical problems are being prosecuted, the majority of them under the direction of A. R. Kemp. One of these relates to the use of rubber and gutta-percha in submarine insulation. The latter is the classical material for this purpose, while rubber has been regarded as an inferior substitute. The supremacy of gutta-percha is due in part to its mechanical characteristic of thermoplasticity, which permits it to be extruded as a continuous insulating layer about a conductor, requiring nothing but the simple process of cooling to convert it into a tough, firm sheath. An even more peculiar virtue of gutta-percha is the stability of its electrical characteristics during prolonged immersion in water. By patient experiments extending over several years it has been demonstrated that the inferiority of rubber in this one respect is due wholly to its non-hydrocarbon constituents. Methods for the elimination of these foreign substances, notably water-soluble salts, proteins, and quebrachitol, which could be applied without damage to the hydrocarbon, have required extended study.

In connection with the water-soluble impurities, an interesting problem in osmosis arose. The absorption of water by rubber is found to be a direct function of its content of water-soluble substances and an in-

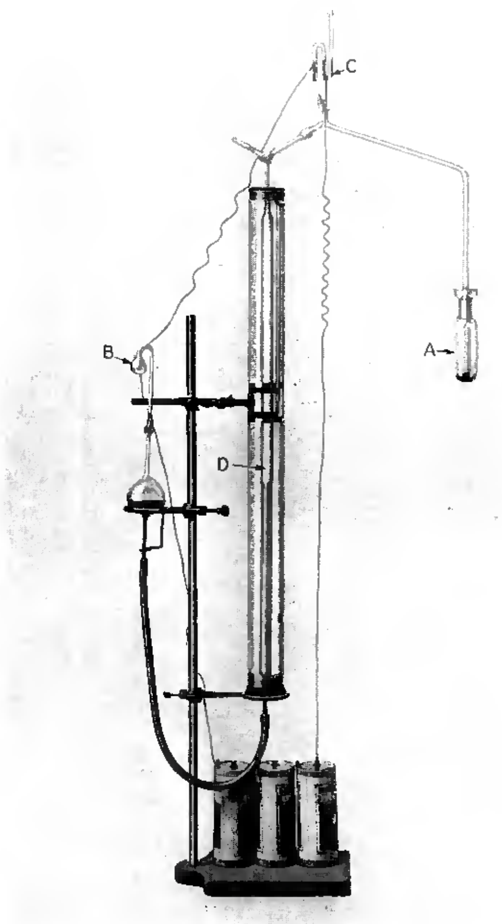


Fig. 1—Apparatus for observing rate of oxygen absorption. Rubber, asphalt, or paint film to be studied is placed in bulb *A* (enclosed in thermostat not shown) and apparatus is filled with pure oxygen. By operation of the manometer *C*, water is electrolyzed in bulb *B*, generating gas automatically in quantity just sufficient to maintain oxygen in buret *D* constantly at atmospheric pressure.

verse function of the salt content of the external water. Thus a sheet of raw rubber which is produced by the evaporation of latex, and therefore contains all the natural water-soluble impurities, will, upon immersion in distilled water, gradually absorb more water than was contained in the original latex, though the sheet still retains its original form, somewhat swollen of course by water absorption. On the other hand, rubber which has been carefully freed of all water-soluble impurities absorbs but a small percentage of water under the same conditions. Fresh-water cables are more liable to degradation from water absorption than cables in salt water, and if the ocean were a saturated solution of sodium chloride the problem of the use of ordinary rubber would be materially simplified.

A development of primary importance in this connection has been the elimination of the proteins from rubber. It appears that the protein constituents form an intricate network which permeates the entire

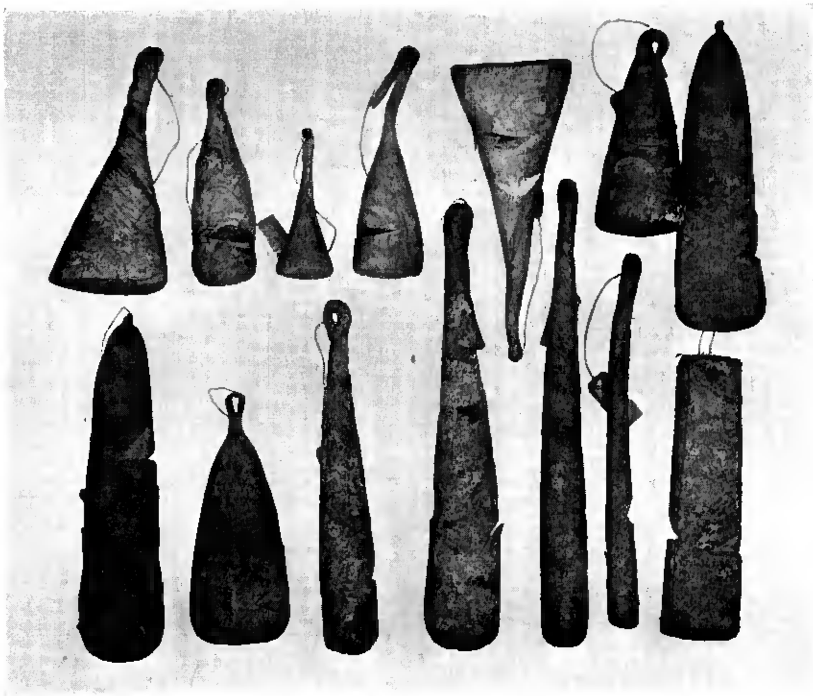


Fig. 2—High-grade gutta-percha in shapes popular with the forest gatherers.

mass. While the amount of proteins present is too small to cause a large absorption of water, they do insure that such water as is absorbed causes a maximum electrical damage by extending the water-bearing filaments through and through the material. The partial hydrolysis of the proteins and the thorough coagulation of any remaining residues by autoclaving the rubber by steam pressure has afforded a simple but effective means of stabilizing the material electrically against the action of water.

Rubber also finds many uses in aerial insulation, a field which is rather backward in its development as compared with the tire industry. The latter has received an incalculable benefit from the large amount of technical research carried on in the last ten years. To a considerable extent it has been our task to adapt technical information derived from the tire industry to the needs of the telephone. The adoption of such expedients as accelerators and anti-aging compounds has contributed a great improvement in the field of insulation. In fact, the aging of rubber in the case of insulation is obviously even more important than in the case of tires. One scarcely expects automobile tires to last for more than a season or two, but rubber insulation must often be exposed to sun and rain for ten years or more.

Rubber for wire insulation must offer resistance to cutting of the wire through the insulation under severe load, such as is produced by a deposit of sleet. Our chemists have therefore been called upon to contribute to the development of a compression-testing machine for rubber which automatically plots the reduction of thickness of wall under an increasing compressive load. By the use of this testing machine surprisingly great variations were discovered and corrected in the material supplied by different manufacturers.

Other useful tools in the study of rubber have been developed. A method has been devised for the direct determination of rubber hydrocarbon in compounds by means of iodine titration. Direct gasometric measurement of oxygen consumption by rubber and other organic materials is performed with a special apparatus for automatically maintaining oxygen pressure constant at one atmosphere. Included in the program are cooperative studies of various accelerated aging tests for rubber, notably the Geer test and that of Bierer and Davis. In these ways we hope we are repaying in part our debt to technologists of the rubber industry, as well as serving our own needs.

Textiles

Another general class of insulating materials in which study has been well rewarded is that of textiles. It has now been quite clearly shown that textile fibers serve as filaments upon which the moisture of the atmosphere is deposited, and that the electrical characteristics of the textiles are determined largely by the thickness and continuity of these water films and the conductivity of the solutions formed by contact with the textiles. The presence of water-soluble impurities in textile insulations is therefore very important and their thorough removal by controlled processes of washing has resulted in vast improvements in both cotton and silk.

For some reason not yet fully understood, silk is much superior to cotton as an insulating material over the usual range of atmospheric humidities. This is true in spite of the fact that cotton absorbs less water than silk at a given humidity, and that, over the usual range of use, silk is more sensitive electrically to a given increment of water than cotton. This is responsible for the extensive use of silk in the electrical industry at several dollars per pound in place of cotton, which may be had for about a tenth as much. The purification of cotton, however, has made it good enough to replace silk for a large number of purposes, especially in telephone cords, and the saving thus effected amounts to several hundred thousand dollars per year for the Bell System. In addition to this practical result, such studies as these have suggested interesting scientific possibilities in the use of electrical measurements for determining the structures of colloids.

Paper

An allied product is paper, which is used in enormous quantities in the construction of the common type of telephone cable. This cable consists of a bundle of wires individually insulated from one another by strips of paper helically served about each. Before being enclosed in a lead sheath the bundles of insulated wires are thoroughly dried and thereafter throughout their use have to be protected from entrance of atmospheric moisture by hermetically sealing the lead covering. The functioning of the cables depends absolutely upon the maintenance of an extremely dry atmosphere within the cable. Our chemists have been called upon for elaborate studies of the effects of minute increments o-moisture upon the insulating qualities of paper and the effect of temf perature upon the electrical characteristics of paper containing various small proportions of moisture. Incident to this task it has been necessary to develop a humidity recorder sensitive to as little as 10 parts per million of water vapor in the air. Such a commercial recorder, produced by Leeds and Northrup at our instance, is in successful use as a guide in controlling the atmosphere of cable-drying ovens. Improved devices for determining the brittleness of cable paper and for judging its predisposition to lose flexibility upon baking have also received attention.

Still another dielectric problem is that of condensers, which are unique in that an insulator of maximum dielectric constant is required, whereas in most electrical apparatus a minimum dielectric constant is sought. For the sake of economy the usual telephone condenser is made of alternate strips of paper and tinfoil which are wound up into a compact roll, and after drying is impregnated with some form of waxy material to bring the capacity to a maximum and prevent subsequent variations with changes of atmospheric humidity. The complexity of the effects of the choice of the impregnating material upon the electrical characteristics of the condenser is most surprising. It might be supposed that condensers of high insulation resistance would also have a high breakdown strength, but this is by no means always the case. It has become evident that the nature of the interface between the individual fibers of the paper and the surrounding waxy material is of vital importance.

Phenol Plastics

A class of materials very widely used in the electrical industry for insulating purposes is well known under the general term of "phenol plastics." Cellulose acetate is another insulating material of excellent electrical characteristics and is superior to the usual fibrous materials owing to its lower water absorption. Although the uses of such materials, both in massive form and as impregnants for fibrous insulation, are very extensive in the telephone field, little chemical work has been done in the Laboratories on them, partly because they are the products of a rather highly developed industry which has conducted a great deal of investigation for us. For certain types of uses, however, the phenol plastics have required some special chemical study from the standpoint of their stability. Being in the nature of condensation products formed by the elimination of water, the phenol plastics are more or less subject to the reversion of the reaction by which they were formed with the production of free phenol and ammonia. For certain uses the presence of these uncondensed constituents or hydrolytic products, as the case may be, is objectionable and has required a special investigation of means of controlling their presence. An important improvement and economy has been effected in certain textile-insulated wires by applying a cellulose acetate lacquer to the exterior so as to partially impregnate the textile. The film of cellulose acetate contributes a continuous smooth surface and a measure of resistance to atmospheric humidity variations.

Conductivity of Copper

Passing from insulators, one naturally thinks of conducting materials, of which of course copper is first in importance. It is a well-known and thoroughly tried principle that the metallic elements in the pure state have a higher electrical conductivity than their alloys. The problem of conductivity of copper is therefore essentially one of purity, and has been fairly satisfactorily solved by the copper-refining industry. We are, however, attacking some special problems in conductivity of copper.

Carbon

A unique conducting material in the field of telephony is carbon. A small mass of granules of this material in every transmitter serves the all-important purpose of converting the variations of the mechanical energy of the voice into equivalent variations of the transmissible electric current. This the carbon does by variation of its electrical resistance with variation of mechanical pressure.

No other material approaches carbon in its usefulness for this purpose, but there are still many obscurities about the functioning of a carbon transmitter. Decades of physical and chemical research have, however, established certain points. Transmitter carbons in general are not highly active carbons in the sense in which we have become familiar with that term in connection with absorptive charcoals. Gas films on the carbon, however, do play some role in their microphonic functioning. Of more practical importance for the present is the evidence that the microphonic effectiveness of carbon is very much dependent on the method of its preparation, and particularly the time, temperature, and atmosphere in which it is roasted. Carbon for transmitters is made from anthracite coal of maximum hardness and low ash content. The roasting processes are designed to produce a material of as high uniformity as possible, especially with reference to the hardness, compactness, and abrasion resistance of the surfaces of the finished product. It has come to be recognized that such physical characteristics of transmitter carbon as these to a great extent determine, not only its original effectiveness, but also its resistance to atmospheric disturbances and its durability under the mechanical and electrical forces exerted upon it during use.

The hydrogen content of carbon has been found a useful index of the time-temperature cycle to which it has been subjected during roasting. As anthracite coal is roasted there is a progressive loss of hydrogen, which, however, does not reach zero value until a temperature above 1500° C. is obtained, at which point the material is converted rapidly to graphite. It is only at upper intermediate temperatures that satisfactory transmitter carbons can be produced.

The study of the hydrogen contents of coal and carbons and of related microphonic behavior has led to a definite theory that there are, contrary to belief of some authorities, only two allotropic forms of carbon—namely, graphite and diamond. The so-called amorphous carbons, according to this view, are complex hydrocarbons in which the carbon greatly predominates over the hydrogen and in which, especially if the carbon has been roasted at a high temperature, the carbon atoms are in part arranged in a graphite lattice.

Vacuum-Tube Filaments

The filaments of vacuum tubes represent a very special form of conductor, the primary function of which is, of course, the emission of electrons. Vacuum tubes are used in repeater sets in all long-distance telephone lines. The filaments in these vacuum tubes consist of platinum alloys coated with the oxides of barium and strontium. It is important that these filaments should be constant in their electrical characteristics and have as long a life as possible. The manufacture of platinum alloys and the methods of coating them have been subjects of study by our metallurgists and chemists for several years, and as a

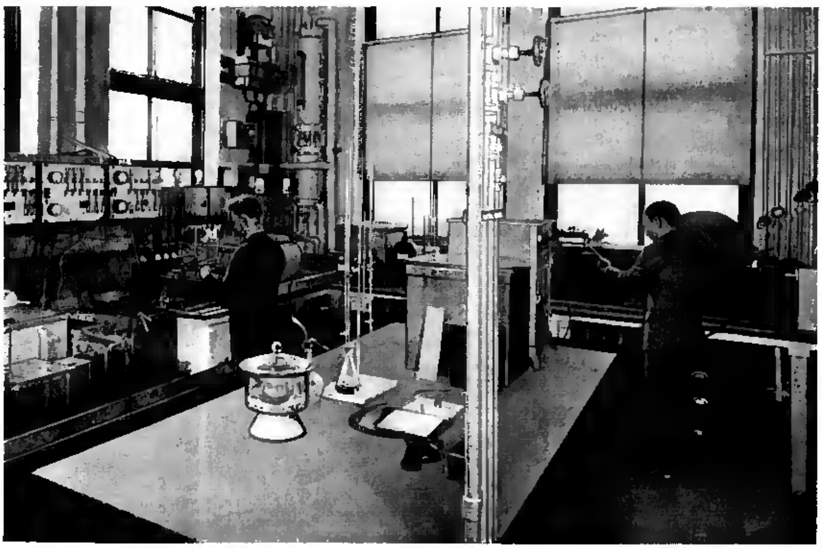


Fig. 3—A corner of the electroplating laboratory.

result many improvements in the manufacture of the filaments have been made. Probably the most outstanding improvement has consisted in the substitution of platinum-nickel and platinum-cobalt alloys for the platinum-iridium-rhodium alloy formerly used as a filament core. This substitution resulted in increasing the life of repeater tubes from a few months to several years.

Metals

The metallurgical group, as far as chemical investigation is concerned, is in charge of J. E. Harris. One of his principal interests is the production in varied forms of special magnetic materials, illustrated by permalloy and perminvar. The field of magnetic materials has been so intimately connected with the fundamental physical theory of ferromagnetism that the primary responsibility for these investigations has been lodged in a physical research group, the members of which have been responsible for fundamental inventions in this connection. The

Fig. 4—Pouring a metallic melt in vacuum by tilting a high-frequency induction furnace.

problems of fabrication and composition control in the experimental work in this connection have, however, afforded much opportunity for the ingenuity of chemists.

A particular case in point is that of the production of brittle forms of such special magnetic materials, the object being to permit the grinding of the metal into a coarse dust. The dust particles are then insulated by deposition of a film on their surfaces and are pressed into rings which form the cores of modified Pupin-type loading coils. This unusual objective of rendering metal brittle has been achieved in general by the principle of introducing an impurity into the melt, which has a tendency to segregate at the grain boundaries, thus facilitating subsequent fracture. Each alloy requires some special consideration, both as to choice of embrittling agent and heat treatment and working schedules for development of proper grain size.

Scarcely any metallic material has given the telephone industry more concern than the lead alloy used for cable sheath. Pure lead is too soft for the purpose and can be too easily damaged mechanically. Years ago about 3 per cent of tin was alloyed with the lead as a hardening agent. Tin was later superseded by 1 per cent of antimony, primarily for reasons of economy, and it has been reported that twenty million dollars have been saved to the telephone system by this substitution alone.

The use of these hardening agents affords an example of the dispersion hardening of metals, which has become familiar to the public most conspicuously in the case of duralumin. It is the belief of metallurgists that the introduction into a molten metal of a constituent, which is precipitated out in very finely divided form upon cooling the metal, diminishes the deformability of the finished material by interposing itself in the slip planes among the atoms of the metal.

Dispersion hardening in a metal as soft as lead represents a rather extreme case, for lead at atmospheric temperatures is approximately as deformable as steel at dull red heat. It has been found that the antimony used as a hardening agent in lead cable sheath tends, especially under the influence of repeated flexings such as those due to thermal expansion, to redissolve in the metal and redeposit elsewhere. In this fashion large particles of the antimony grow at the expense of small particles, and the hard-worked portion of the metal is eventually deprived of antimony content and fracture occurs.

This has been the source of considerable trouble in aerial cables, especially at the bends in the cable which occur at the poles due to expansion. The working out of this problem in fatigue of metals has been a long process, but promises to bear further fruit in the development of better hardening agents. One of these, a joint development with the Western Electric Company and one which still remains to be tested on a commercial scale, is calcium, which in the minute proportion of 0.04 per cent has been found in laboratory experiments to produce a hardening well surpassing that of 1 per cent of antimony.

Another interesting metallurgical problem is that of solders for use in wiping joints in telephone cables. Somewhat to our surprise we found that some of the supposed prejudices of workmen responsible for cable splicing were well founded and that it is scarcely practicable to make a satisfactory wiped joint with a lead-tin solder containing less than 38 per cent or more than 42 per cent of tin. New solders may well grow out of our study of why and how the old-fashioned solder works.

The interests of the telephone system extend of course to many other metallic materials, notably iron and steel, brasses and bronzes, die-casting alloys, etc. For the most part, however, progress in these fields has been along the lines of that of other industries. Part of our metallurgic shops are largely devoted to the melting, casting, and fabrication of a great variety of alloys into wire rods or sheets of specified dimensions for experimental trials in electrical apparatus design.

Wood and Its Preservation

It is a far cry from metals to wood, and particularly to wood preservation, which is one of our important chemical interests. The telephone pole is our most urgent concern. Large numbers of poles of cedar, chestnut, and southern pine are in use and the greater part have been subjected to a preservative treatment. Pine poles possess a layer of 2 or 3 inches of sapwood which is subject to impregnation under pressure, and such a treatment with creosote has long been the standard practice in the system.

The problems in this field are innumerable. We will mention only a few. Given a train load of telephone poles, how does one determine the average quantity of creosote they contain and the uniformity of distribution from pole to pole and in various parts of a given pole? No two trees grow alike. Soil, climate, sun exposure, accidents in past histories such as fires in forests—all make for peculiarities of growth in each individual tree. These peculiarities reflect themselves in the absorption of creosote, so that it is entirely possible for two poles treated simultaneously in the same cylinder to differ by a factor of 5 or even 10 in the over-all creosote content per unit volume. How can one obtain a sample which will be representative of a large group of such poles?

The sampling problem was approximately solved by taking a sufficient number of cylindrical solid cores with an increment borer and splitting these borings diagonally along a length which represents the approximate radius of the pole. Mathematically such a tapered cylinder approaches a wedge such as would represent a true sample of the pole's cross section. The borer holes are plugged with a creosoted peg and the poles are still fit for use, so one can sample as many poles as he likes.

This method is gradually being applied to a study of the content and distribution of creosote, not only in new poles, but also in those that have been in service for varying periods of years. In order to determine how much creosote is present and also its present wood-preservative value, we have adopted a biological method of testing the toxicity to pure fungus cultures of creosote extracts from such old wood. An interesting story could be written about the specific resistance of various species of wood-destroying fungi to the numerous toxic agents that have been used and proposed.

In passing upon the merits of a new preservative it is difficult to predict its permanence in the wood when exposed to the weather. A promising method, though one which does not offer the ultimate in economy of time, is the use of small twigs or saplings which are impregnated with the preservative in question and exposed in groups to the action of the weather in a fast rotting climate.

In this test we are depending for acceleration entirely upon the reduction of the dimensions of the wood, while preserving at least the more salient features of wood structure by using natural twigs or stems rather than artificially shaped pieces. In applying this principle of reduction of dimensions for purposes of acceleration, we took a leaf from our book of experiments on submarine insulation in which we had found a strong case of parallelism between the absorption of water by rubber and the saturation of a material by heat as epitomized in Fick's law. The time required to reach a given degree of saturation is approximately inversely proportional to the square of the thickness of the specimen. The depletion of creosote from wood appears to be an inverse process, with of course some complications. Reasoning by analogy we hope by the use of small stems to shorten the time required for depletion of creosote and other preservatives and the beginning of rot in wood by ten to thirty years—a great economy in patience.

Electrochemical Investigations

Another large field of chemical investigation comes under the general head of electrochemistry, which for convenience includes corrosion, corrosion prevention, and finishes of both metallic and organic (paints, etc.) types. This work as well as the analytical laboratories, is under the supervision of R. M. Burns.

Among the electroplating developments has been the successful deposition of permalloy from a bath containing iron and nickel salts. The composition of the alloy containing 79 per cent nickel and 21 per cent iron can be maintained constant to less than 0.5 per cent. It is of interest to find that the alloy is deposited from the bath as a solid solution and that it has desirable magnetic properties.

The corrosion work consists of both corrosion testing and fundamental studies on the mechanism of corrosion processes. The corrosion tests are carried on under normal service conditions and by laboratory-accelerated processes. The results furnish guidance in engineering decisions as to the use of materials. More fundamental investigations in the field of corrosion have to do mainly with the study of the electrochemistry of corrosion reactions, film formation, etc. It is of importance, for instance, in the development of corrosion theory to determine

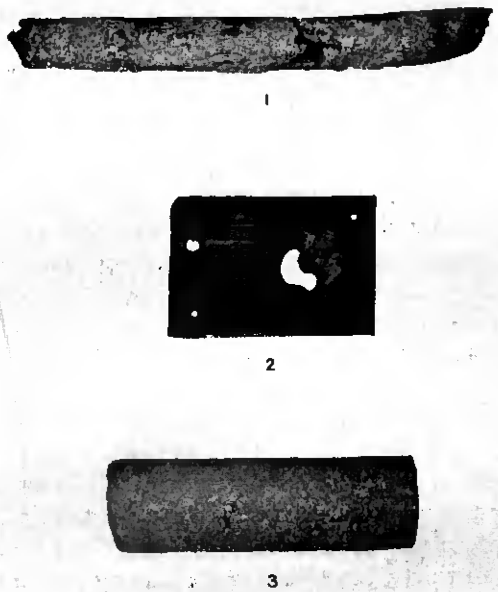


Fig. 5—Corrosion of lead cable sheath. (1) By direct contact with a pernicious soil. (2) Stray current anodic corrosion. (3) Acetic acid corrosion in fir duct.

the effect of the environment, and of the passage of small electrical currents, upon the anodic and cathodic behavior of pure metals, and work of this nature is being carried on.

Underground corrosion is an important part of the field. The most striking examples of this phenomenon are found in steel and iron structures and particularly in the lead sheaths of subterranean cables. Stray currents from trolley systems often cause such corrosion and an elaborate and expensive electrical bonding system is maintained in order to minimize these troubles. The physical and chemical nature of the soils and the underground waters and atmosphere often play an important role in determining the kind and extent of corrosion by stray current.

Other occasional cases of electrochemical corrosion have been encountered in which stray current, though present, does not arise from trolley-line power houses. In one large city it was found that a battery covering a square mile or more of area had been inadvertently created, such that it affected a large part of the cable system in the center of the city. The cinder fills underlying the duct runs in this area contained enough carbon to serve as one electrode, while the iron-pipe systems supplying gas and water to the city furnished the other electrode. The

Fig. 6—Experimental metallurgical shop.

moist soil afforded a conducting path for a galvanic current that wrought a widespread damage to telephone cables in the area.

In another and much larger area widespread injury to cables came about through the presence of traces of acetic acid in the air in wooden duct systems. The source of this acid was the wood itself, which happened to be of a rather highly acid variety. The natural acidity of the wood was further increased by the somewhat drastic process of heating which was necessary to secure a fair penetration of the wood with creosote. In so far as the creosote penetrated, the acid produced in heating was neutralized to a great extent by the nitrogenous bases in the creosote. Often, however, the total acid produced far exceeded the neutralizing power of the creosote bases contained in the external shell of creosoted wood. The sheath of many miles of cables underwent a partial conversion into white lead via the classical Dutch process which, though highly regarded by paint manufacturers, became anathema to telephone engineers. The difficulty was met by fumigating the ducts in service with a dilute ammonia-air mixture and by choosing a less acid and more easily treatable wood for future construction.

Underground corrosion of other metals, notably of iron and steel, is also often serious. In the alkali soils of the southwest anchor rods for telephone poles have sometimes corroded through in a few months. Marshes represent another severe exposure for iron and steel, as, for example, in the form of loading coil cases. A newly introduced form of telephone cable for direct burial in the soil demands careful consideration from this standpoint. A variety of protective finishes, chiefly of asphaltic or pitchy nature, have been studied in this connection. Some remarkable cases have been noted, in which a finish that proved to have a superior protective effect in one highly corrosive soil was worse than useless in another soil which had been regarded as less corrosive in the general sense.

The chemistry and physics of soils from many areas have required attention with the control of corrosion as an object. Particle size, saline content, and composition of subsoil atmospheres each has an influence.

In a like way the telephone chemist must concern himself with atmospheric causes of corrosion in equipment above ground, especially in central offices. Moisture and dust contribute to electrical leakage from point to point through the complicated assemblies of electrical equipment. Corrosion products of such leakage may build up at critical points and interfere with contacts, or essential though usually minute portions of equipment may be etched away. Even faint tarnishes on metallic contacts can so increase contact resistance as to imperil signaling. In industrial areas soot and traces of sulfurous gases add materially to these hazards.

Finishes

It is partly to avoid such difficulties and partly for the equally utilitarian purpose of a good appearance that metal telephone apparatus receives some special form of finish coating, whether paint, varnish, lacquer, or electroplated surfacing. In the selection of such finishes we have drawn heavily upon the scientific work of our confrères in allied fields, but have still found ourselves faced with peculiar difficulties.

A great deal of the truly excellent scientific work on finishes has been done by manufacturers with the idea of disclosing uses which will justify the sale of a particular material. But comparative data on the durability of very dissimilar finishes as, for example, galvanized coatings in contrast to cellulose lacquers, are usually lacking. The fact that finishes are often used primarily for decorative purposes on relatively short lived articles has limited the study of the durability of such coatings. This is reflected in the fact that nearly all the scientific work in this field refers to outdoor exposures where corrosion tends to occur rapidly. Indoor exposures are commonly regarded as so mild as to be negligible. Changing fashions, as in the case of furniture, often bring an obsolescence so early as to be prohibitive were similar consideration to be applied to telephone plant. The prevalent custom of trading in one's motor car for a new model each year is a factor in another large industry involving extensive use of finishes, which tends to put great emphasis upon initial beauty rather than permanence over periods of ten to twenty years, such as must be considered in telephone plant.

With the manufacturing and operating telephone companies tied into a single system, in which the Bell Laboratories' responsibility is to insure quality of product, there can be no unloading of defective apparatus upon the consumer, for the manufacturer is liable in the last analysis for defects which may appear only after years of use.

These considerations have received special emphasis in the discussion of finishes because they afford an excellent illustration of the point. The same sort of considerations, however, apply to nearly all the problems with which we are concerned, to such an extent that our chemical staff is widely thought of in our own organization as a group of specialists in the "permanence" of materials. Such an emphasis by the general management upon ultimate economy rather than first cost alone would doubtless be welcomed everywhere by thoughtful technical men throughout the country. It is a source of peculiar pride to the staff of the Bell Laboratories that the nature and organization of their business is such as not only to permit such an attitude but also aggressively to promote it.

The Trend in the Design of Telephone Transmitters and Receivers [1]

By W. H. MARTIN and W. F. DAVIDSON [2]

This is a report of the Joint Subcommittee on Development and Research, National Electric Light Association and Bell Telephone System. It was prepared by the Chairmen, respectively for the Bell System and the N.E.L.A., of the Project Committee assigned to this study.

The report reviews the history and present trend of the design of telephone transmitters and receivers, particularly from the standpoint of their response frequency characteristics, and discusses the possibility of obtaining a reduction in the effect of line noise by shifting their points of maximum response. It is concluded that no advantage from this standpoint is indicated inasmuch as it has been found that the distribution with frequency of the extraneous energy on telephone toll lines is approximately uniform over the more important portion of the frequency range. It is further stated that the present trend in improvement of the response characteristics of transmitters and receivers is in the direction of reducing the difference between their maximum and average response.

IN the beginnings of the telephone, the outstanding marvel was that the devices used as transmitters and receivers could perform the necessary conversions between speech sound waves and electrical waves. In the application of these devices, however, it was early appreciated that the range and cost of telephone circuits were directly

[1] Editor's Note: In this issue of the *Bell System Technical Journal* there are two papers and one report dealing with various phases of the inductive coordination problem, which have had their origin in the work of the Joint Subcommittee on Development and Research of the National Electric Light Association and the Bell Telephone System.

This organization is one of the subcommittees of the Joint General Committee of the N.E.L.A. and Bell Telephone System, which has for its general objective the working out of methods of procedure whereby problems involving the physical relations between the plants of the electric supply companies and the telephone companies may be handled cooperatively on mutually satisfactory bases. The questions involved are largely of an engineering character, and to carry on that phase of the work the Engineering Subcommittee of the Joint General Committee was appointed. The Engineering Subcommittee has recommended certain broad principles of cooperation as well as the adoption of more detailed principles and practices, which were accepted by the Joint General Committee and published in 1922.

As a result of further recommendations by the Engineering Subcommittee the Joint Subcommittee on Development and Research was organized. It is charged with the conduct of technical investigations, the accumulation of data, and the development of engineering methods for use in the solution of problems of coordination. Its work is organized under a number of subordinate committees known as "Project Committees," each of which is assigned a certain range of subjects for study.

The first volume of Engineering Reports of the Joint Subcommittee on Development and Research, containing a considerable part of the technical information thus far developed by the subcommittee has recently been published (April, 1930).

[2] N. E. L. A. Bulletin, Aug., 1930.

dependent upon the efficiency with which these instruments made these conversions. Experimental activities were, therefore, soon directed to increasing the efficiency of these instruments and especially to getting a more efficient transmitter than the forerunner of the present telephone receiver which initially was used both as a transmitter and a receiver. The outcome was the carbon contact transmitter which provided a means for drawing upon an outside source of energy in the process of converting sound waves into electrical waves and thus combined in the transmitter the function of a converter of energy with that of an amplifier.

Since that time numerous important improvements have been effected in both the carbon transmitter and the magnetic receiver but the general principles of both are still employed today in the best practical instruments for commercial telephony. Both of these instruments employ vibrating diaphragms which, like other mechanical vibrating systems, have regions of maximum response due to resonance between the mass and elasticity of the diaphragm. With these resonant effects inherent in the structure, it was natural to place them in the frequency range so as to obtain the maximum benefit. In accordance with this, the reproductions of speech sounds obtained with these resonances located at different points were listened to and the judgment reached, taking into account both the intelligibility and naturalness of the reproduced sounds, that they should be placed around 1,000 cycles. It was found that a material shift in the point of maximum response of the circuit to a lower value made the output sounds "boomy" and to a higher value rendered them "thin." While precise means for measuring the effects were not available at that time, subsequent work has substantiated this choice as a wise one. Investigations of the frequency components of speech sounds have shown that the principal components of about half the vowel sounds lie below 1,000 cycles and of the other half are about equally divided above and below this point. Articulation tests have demonstrated that the frequency range which covers about an octave each side of 1,000 cycles, namely from about 500 to 2,000 cycles, includes the more important frequency components in speech from the standpoint of intelligibility. The frequencies below this range are important primarily for naturalness and those above for intelligibility and also for naturalness. The location of the region of maximum response of the telephone circuit in the neighborhood of 1,000 cycles emphasizes then this 500 to 2,000-cycle range and meets well the requirements of both intelligibility and naturalness.

In addition to the diaphragm resonances there are also inherent resonances in the enclosed cavities which are associated with these diaphragms, such, for example, as the mouthpiece of the transmitter and the cases in which the transmitter and receiver units are placed. In the present type of deskstand transmitter the several resonances are so located as to give a fairly broad maximum response in the range between 1,000 and 2,000 cycles and the resonance of the receiver has been placed so that its maximum response is around 1,000 cycles. It is seen then that the inherent resonances of these instruments have been located in the more important part of the voice-frequency range and have been utilized to increase their efficiencies in that range.

The remarkable performance of the granular carbon type of transmitter merits some indication of its accomplishment. Its conversion of the complex speech waves into equivalent electrical waves has been improved from time to time and now the most efficient type of transmitter which is in general use, when energized with the direct current which it gets on short loops, has an electrical output which is more than a thousand times the magnitude of the acoustical power which is delivered by the speaker. Furthermore, it provides this conversion and large amplification at a low cost. Since the average energy given out by a speaker in carrying on a telephone conversation is of the order of 10 microwatts, the large stepup in power from the acoustic waves entering the transmitter to the electrical waves leaving it is of vital importance in affording telephone service at a reasonable cost and also in rendering the telephone system less susceptible to the effects of interference currents.

While large improvements have been made in the receiver, the efficiency of the present instrument is very low in comparison with many other types of energy converters which it is considered practicable to use. For the receiver, the average ratio of the acoustic power output to the electrical power input is below 1 per cent. It is possible to increase materially the efficiency of the receiver used in commercial telephony but this would bring up the noises on the telephone circuit. Also there are limitations upon the maximum efficiency of the combination of transmitter and receiver due to crosstalk between telephone circuits and to the fact that with loud talkers over short telephone connections the combination of present instruments is close to the point of giving uncomfortably loud sounds in the ear of the listener.

In considering the performance of the transmitter in the plant, it is customary for many reasons, important among which is the battery supply circuit, to take the combination of a transmitter, a station set, a typical loop connecting the set to the central office and the cord circuit from which is supplied the direct current for energizing the transmitter. Likewise, the receiving system of the circuit may be considered to consist of the cord circuit, the loop, the set and the receiver. For the connection of such transmitting and receiving systems through a distortionless trunk, the response characteristic of the overall circuit, giving the relation between the power delivered by the receiver and the power available at the transmitter, shows a variation of about 30 db in the range from 500 to 2,000 cycles with the maximum response slightly above 1,000 cycles. This characteristic applies to the type of deskstand apparatus which is now the most generally used station equipment in the Bell System.

With the development of the telephone art numerous ideas have naturally been investigated for improving the performance of the transmitter and the receiver. Taking into account the various considerations and possibilities, the present procedure is on the basis that further improvements should come from reduction in distortion rather than from increases in the maximum response. For practical instruments the desire is primarily to reduce the distortion without sacrificing the average efficiency over the important part of the voice range. Means have been developed for reducing the distortion in these instruments but in general such improvements have involved material reductions in efficiency. For example, a very high degree of freedom from distortion is realized in the type of carbon transmitter which has been so widely used for pickup work in radio broadcasting. This transmitter, however, requires a powerful amplifier to bring its output to a value comparable with the type of transmitter used in commercial telephony. It has been possible also to obtain large reductions in the distortion of the receiver but here, too, large sacrifices in efficiency have attended this accomplishment. Material progress has been made, however, toward the ideal of a combination of low distortion without sacrifice in efficiency.

Some of these improvements have been incorporated in the transmitter which is used in the handset type of station apparatus. The frequency response characteristic of transmitting and receiving systems such as described above, but using the handset instruments instead of the deskstand, shows a variation of only about 20 db in the range from 500 to 2,000 cycles, and with the handset instruments this same variation of 20 db covers the range from 500 to 3,000 cycles. The handset transmitter presents, therefore, a material advance from the standpoint of reducing distortion.

The proposal has been made at various times that the interference situation might be helped by providing still more efficient transmitters. The various possible means of still further increasing the transmitter efficiency, however, are attended by many difficulties and complications. The efficiency at the point of maximum response could, of course, be increased by piling up the several resonances which have been referred to, but this would give serious distortion effects which would more than offset any increase in loudness which was obtained. In general, it has been found that any improvements which might permit higher maximum responses can be utilized to give greater benefit in the reduction of distortion. Moreover, any large increase in the transmitter efficiency would require measures such as the reduction of the efficiencies of the receivers in order to avoid increased crosstalk effects between circuits and uncomfortably loud transmission over short connections.

When the program of the joint development and research work of the N. E. L. A. and Bell System was formulated there was the idea that in view of the resonant characteristic of the telephone receiver, some benefit might be obtained in the performance of the telephone circuits in the presence of line noise by shifting the point of maximum response of this instrument. Some cases had arisen where pronounced harmonics in the power system in the neighborhood of 1,000 cycles caused serious troubles in nearby telephone circuits and it was felt that if this condition were found to be prevalent in power circuits, some relief in the interference situation might be obtained by shifting the point of maximum response away from this region. The investigations which have been carried out under Project 4 of the N. E. L. A.-Bell System Joint Development and Research Subcommittee, of the noise on telephone lines in different parts of the country have shown that the average distribution of energy with frequency is approximately uniform over the range from 300 to 2,000 cycles with, however, a pronounced dip in the region around 1,000 cycles. A similar decrease in the energy of the components around 1,000 cycles is also shown by the results of the investigations made under Project 5 on the wave shapes of electrical power machinery. With this situation, shifting the maximum response of the telephone receiver away from its present location would thus in the average case be placing it in a region in which larger amounts of interfering currents are to be found. Moreover, examination of the data showing the distribution with frequency of noise currents, indicates that on particular circuits this distribution is by no means uniform but in many cases is materially higher in the region below 1,000 cycles and in other cases materially higher in the upper regions. It would not appear, therefore, that a shift in the maximum response of the telephone receiver would on the average give an improvement from the interference standpoint.

Furthermore, any compromise in the instrument characteristics to favor the line noise conditions, should not have an adverse effect on the many connections on which noise from power systems is unimportant. As has been noted, a material shift in this maximum response would have a marked effect on the naturalness of the reproduced sounds. On the whole, then, no advantage to the interference situation has yet been indicated for shifting the resonance of the receiver.

In accordance with these considerations, the present effort in the development of telephone transmitters and receivers is being directed along the lines of reducing the deviation between their maximum and average responses. Any improvements in the instruments which it may be found practicable to make will be in the direction of increasing the intelligibility and naturalness of the telephone conversations, and the justification for their adoption will include their effectiveness in the presence of typical distributions of interfering currents on the telephone lines.

Mutual Impedances of Ground-Return Circuits

Some Experimental Studies *

By A. E. BOWEN and C. L. GILKESON

This paper describes some of the results of the work of the Joint Development and Research Subcommittee of the National Electric Light Association and Bell Telephone System on the mutual impedances of ground-return circuits.

The first part of the paper deals with some experiments which were performed to establish an experimental background for the testing of theoretical ideas. Different theories, one involving an "equivalent ground-plane," a second a d. c. distribution in the earth, and a third an a. c. distribution in the earth, are discussed in the light of the experimental results. While none of these is adequate to explain all the observed phenomena, each has a field in which it can be made useful.

The second part of the paper is devoted to a description of practical means for predetermining the mutual impedances of power and telephone lines. This involves an experimental determination of a curve of mutual impedance as a function of separation in the region of the proposed exposure and the calculation of the overall mutual impedance between the proposed lines from this curve and the dimensions of the exposure. The results of trials of this method in two locations are given which indicate that itshould be of sufficient accuracy for engineering purposes.

Introduction

THE magnitude of the inductive coupling between power and telephone lines is a factor of fundamental importance in problems of coordination to prevent interference between these two classes of lines. Accordingly, this is one of the subjects under investigation by project committees of the Joint Committee on Development and Research of the National Electric Light Association and the Bell Telephone System. It is the purpose of this paper to present the results of some work which has been done under the auspices of the Committee on one phase of this problem, namely, the mutual impedance of ground-return circuits.

The mutual impedance of two ground-return circuits is determined by measuring the ground-return current in one circuit (the "disturbing" circuit) and the open-circuit voltage at the terminals of the second circuit (the "disturbed" circuit). The vector ratio of the open-circuit voltage to the ground-return current is then defined as the mutual impedance of the two circuits.

For any normal or abnormal operating condition of a power system,

* Presented at the Summer Convention of the A. I. E. E., Toronto, Ontario, Canada, June 23–27, 1930.

the currents, either at fundamental frequency or at any harmonic frequency, in any of the lines can be resolved into components, some of which are entirely confined to the wires while another component flows in a circuit composed of all the wires as one side with the ground as a return path. The work which is described in this paper deals with the magnitude of the induced voltages on exposed telephone lines caused by the latter component. It has been directed to two ends, first, the establishment of an experimental basis for the study of the physical factors involved in the inductive coupling of ground-return circuits, and second, the development of practical methods to enable the advance calculation of the mutual impedances of power and telephone lines. The work is accordingly presented in two parts; first are given the results of tests made at a field laboratory in which testing conditions could be controlled, and second, tests in which the practical side of the problem was investigated are described.

Cross Keys Tests and Theoretical Background

Cross Keys Tests. An extended series of measurements was made at a field laboratory operated by the subcommittee near Cross Keys, New Jersey, about 20 miles southeast of Camden. A single conductor, located about 34 ft. above the ground and 8500 ft. in length, was available for the disturbing circuit. For disturbed circuits, 500-ft. lengths of insulated wire were laid on the earth parallel to the disturbing conductor, at several separations, as shown on Fig. 1. Grounds were pro-

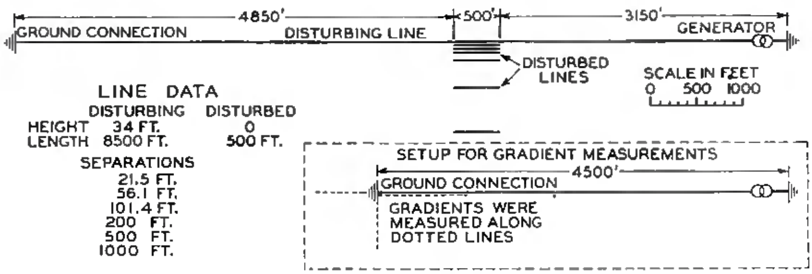


Fig. 1—Cross Keys tests—experimental setup.

vided at the ends of each line. Ground-return current was transmitted over the disturbing line at 60 cycles from a commerical source, or at frequencies between 100 and 1000 cycles from a vacuum-tube oscillator with power amplifier. The measuring instrument was an a.-c. potentiometer, equipped with suitable filters so that the observations were unaffected by the presence of harmonics in the disturbing current.

At several frequencies within the range from 60 to 1000 cycles the

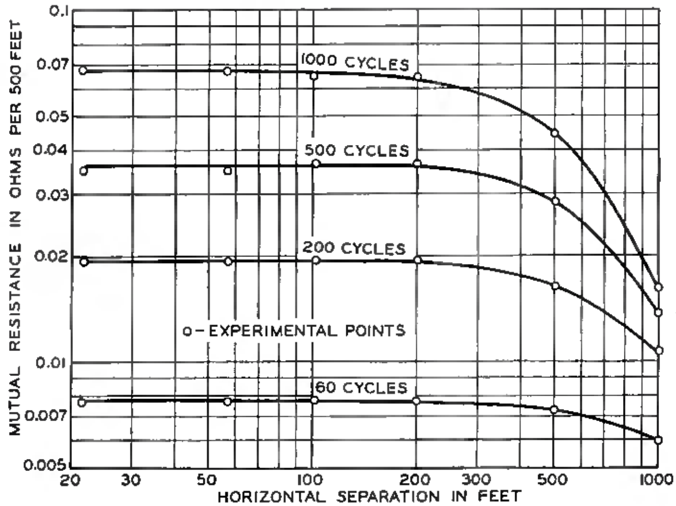


Fig. 2—Cross Keys tests—mutual resistance.

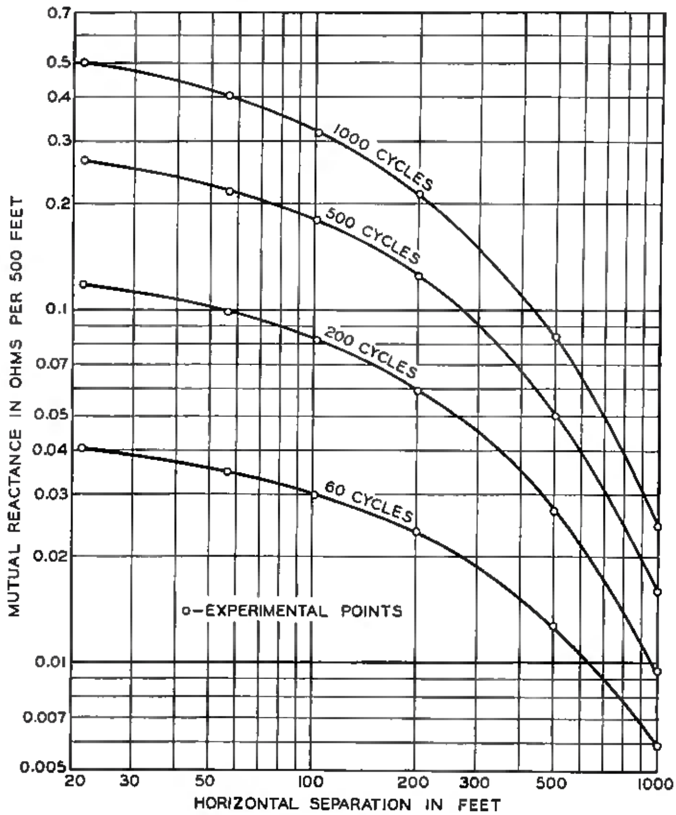


Fig. 3—Cross Keys tests—mutual reactance.

current in the disturbing line, the open circuit induced voltage in each of the short ground-return circuits, and the phase angle between these two quantities were measured. The mutual impedances were derived from the ratio of the induced voltage to the inducing current, in accordance with the definition. The results of these tests are given on Figs. 2, 3, and 4. Fig. 2 shows the resistance components, Fig. 3 the reactance components, and Fig. 4 the magnitudes of the mutual impedances.

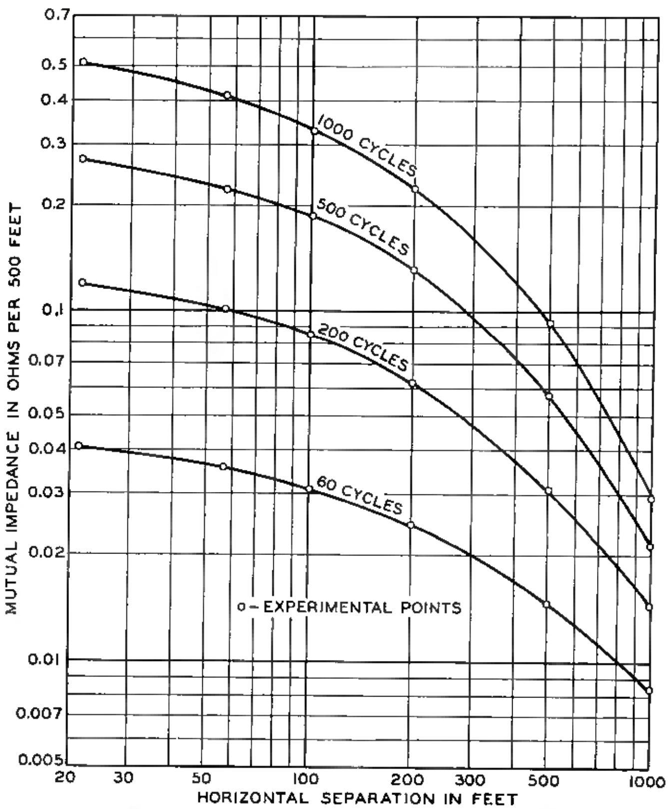


Fig. 4—Cross Keys tests—mutual impedance.

The measurements described above were made with the object of investigating the mutual impedances of ground-return circuits in which the ground connections on the disturbing line are sufficiently removed from those on the disturbed circuit so that effects due to proximity of the grounds may be ignored. The results presented above were supplemented by observations demonstrating that the induced voltage in a

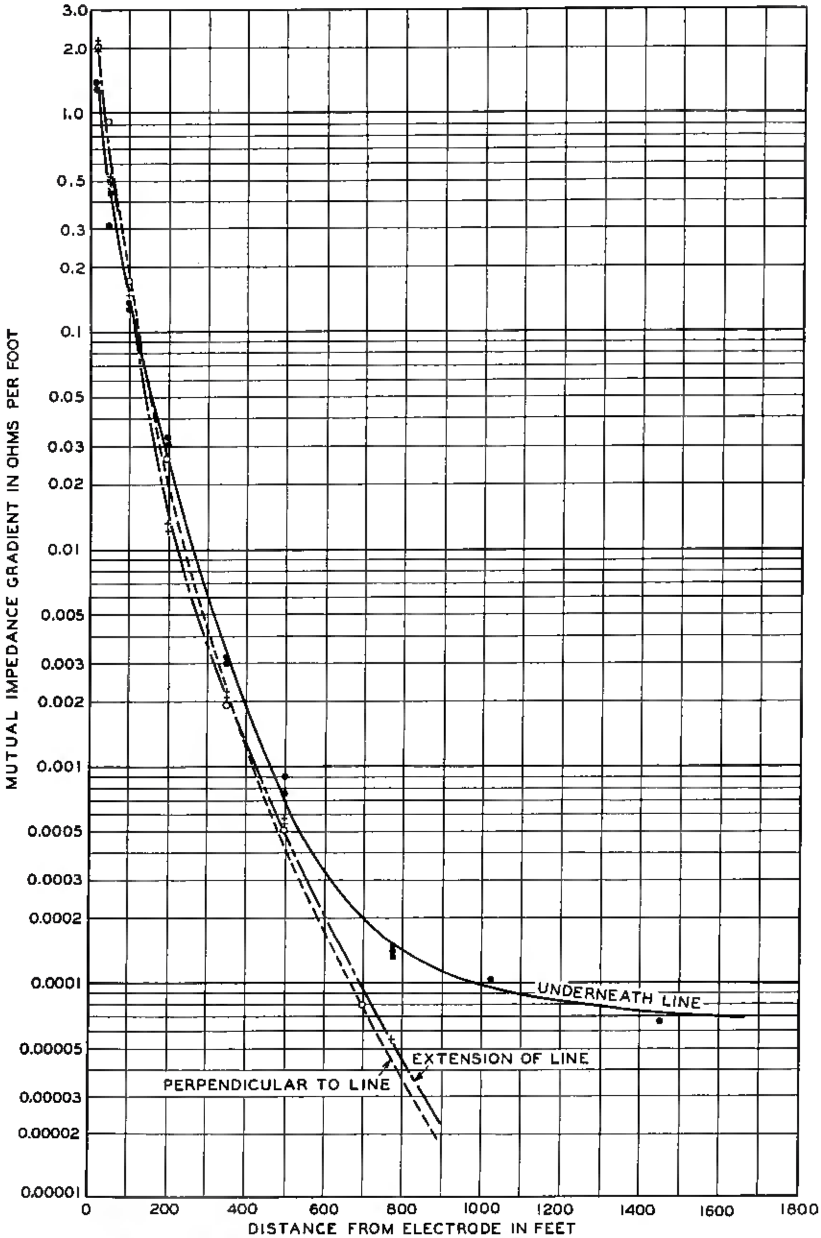


Fig. 5—Cross Keys tests—mutual impedance gradient in vicinity of grounding electrode. Experimental curves.

parallel circuit was closely proportional to the length of the circuit and that the voltage induced in a ground-return circuit extending perpendicular to the disturbing line was exceedingly small.

As the points of grounding on the disturbed circuits approached those of the disturbing circuit this proportionality no longer existed nor was the voltage in a perpendicular circuit of negligible magnitude. A second series of tests was therefore conducted to determine the nature of this effect and the area in which it was of importance.

In these tests voltages were measured in very short disturbed circuits extended along radii converging on the ground electrodes on the disturbing line. At each location the circuit was made progressively shorter until the quantity measured per unit length was practically independent of the length. Thus the gradient of the mutual impedance, in the direction of the radius at the point of measurement was determined. These measurements were made only at a frequency of 60 cycles.

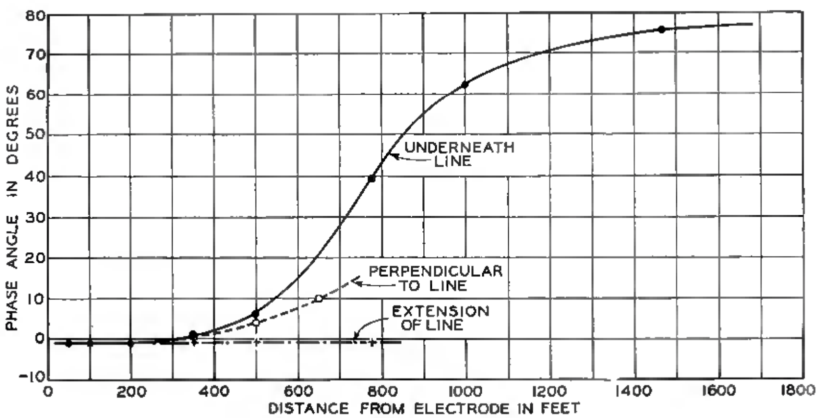


Fig. 6—Cross Keys tests. Phase angle of mutual impedance gradient in vicinity of grounding electrode. Experimental curves.

The resulting observations are shown in Figs. 5 and 6. Fig. 5 shows the magnitude of the mutual impedance gradient along three radii, one radius being directly under the disturbing line, the second perpendicular to it, and the third along the extension of the line. Fig. 6 shows the corresponding phase angles. Under the disturbing line, as the distance from the grounding point is increased, the gradient approaches a constant value and the phase angle changes rapidly from a very small value to an angle approaching 80 degrees. Along the latter two radii, however, the magnitude of the gradient appears to decrease indefinitely and the phase angles are smaller.

A more complete analysis of the results of both groups of tests is given in connection with the discussion of theory which follows:

Equivalent Ground-Plane Theory. The equivalent ground-plane method of computing the mutual impedances of ground-return circuits utilizes a very simple formula and has been in common use for a number of years. A derivation and discussion of the formula together with some experimental results are given in the report published by the California Railroad Commission in 1919.[1]

This method assumes that the returning earth current may be considered as flowing in a hypothetical plane surface of perfect conductivity located some distance below the actual surface of the earth. This surface is usually termed the "equivalent ground-plane." The depth of the equivalent ground-plane below the actual surface of the earth varies in different locations from about 50 ft. to 5000 ft. or more, depending upon the character and resistivity of the earth and the frequency.

This method is subject to the objection that it fails to represent completely the observed phenomena. For instance, the method represents the mutual impedance only with a reactive term, while the experimental results indicated a substantial resistance component, particularly at the wider separations and higher frequencies. Furthermore, no attempt is made to explain the phenomena observed in the neighborhood of the ground electrodes. However, in one respect the theory leads to results comparable to those observed; the magnitudes of the mutual impedances as observed under conditions in which end effects are negligible can be checked reasonably well with a suitable choice of ground-plane. Comparisons demonstrating this point are made on Fig. 7, where it will be seen that the curve of experimental mutual impedance for a frequency of 60 cycles can be fitted very well by a calculated curve with a ground-plane depth of 835 ft. That the depth of the equivalent ground-plane depends on the frequency is seen from the fact that to fit the experimental curve at 500 cycles requires the use of a ground-plane depth of 385 ft.

Method Assuming D.-C. Distribution in the Earth. For an earth of uniform conductivity, the distribution of the current in the earth for a ground-return circuit energized from a d.-c. source has been employed by G. A. Campbell[2] to derive formulas for the mutual resistance and inductance of ground-return circuits. The mutual resistance is expressed by a very simple formula which involves only the earth resistivity and the distances between the points of ground connection on the

[1] See Bibliography.

[2] See bibliography.

disturbing and disturbed circuits. For the calculation of the mutual inductance formulas and graphs requiring only a knowledge of the mutual arrangement of the wire parts of the disturbing and disturbed circuits with respect to each other and the earth are given. The mutual inductance is independent of earth resistivity. While these formulas are, of course, strictly applicable only for direct currents, it is to

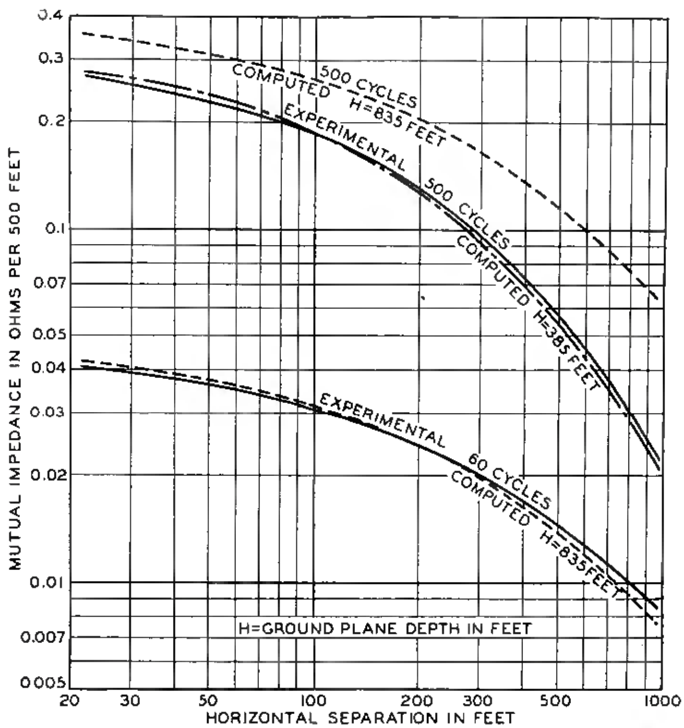


Fig. 7—Cross Keys tests. Ground plane theory. Comparison of measured and calculated values of mutual impedance.

be expected that at sufficiently low frequencies the ground-current distribution would not differ appreciably from that for direct current, and hence for these frequencies, these calculated d.-c. mutual resistance and inductance should approximate the actual values. In the paper referred to, some experimental results at frequencies of 25 and 60 cycles supporting this point of view are presented.

The experimental curves of Fig. 2, which were obtained from measurements at Cross Keys on the 500-ft. disturbed lines near the middle of the 8500-ft. disturbing line, indicate a pronounced increase in mutual resistance with increase in frequency in the range from 60 to 1000 cycles. These results have been replotted on Fig. 8, and it is apparent that for separations within the range of 20 to 500 ft. the mutual resistance increases rapidly in almost linear relation to the frequency. For the frequency range and circuit lengths involved in this series of tests, it would appear that a formula for the mutual resistance, based on a d.-c. distribution in the earth is inadequate.

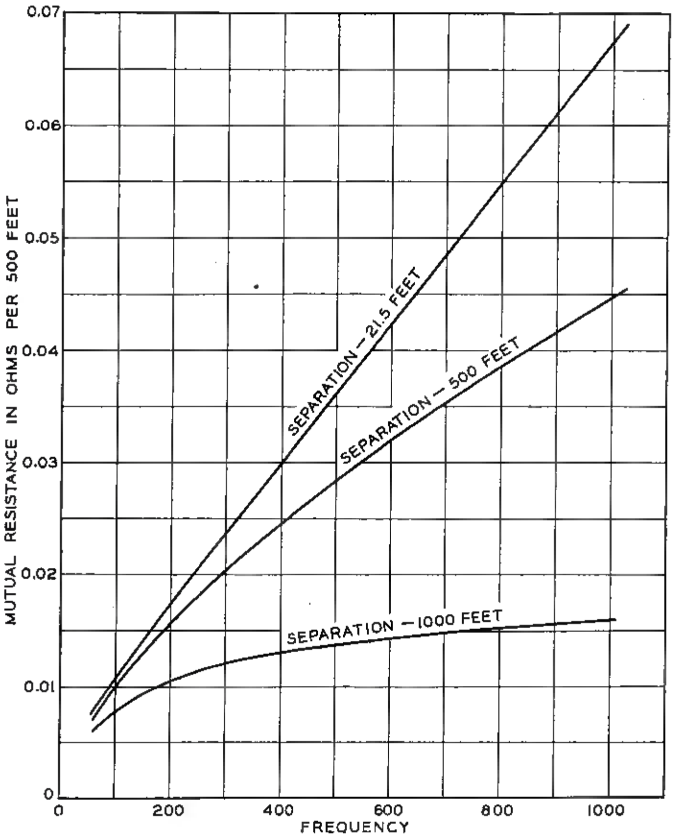


Fig. 8—Cross Keys tests. Variation in mutual resistance with frequency and separation.

The mutual-inductance curve of Fig. 9 has been computed according to the formulas given by Campbell, and for comparison purposes the mutual inductances derived from the mutual reactances shown on Fig. 3 are also plotted. It will be seen that the observed mutual inductances decrease as the frequency is increased, and that while the trend of the observed values is towards agreement with the calculated values as the frequency is decreased, the agreement is far from good at 60 cycles, the lowest frequency used in these tests.

In the immediate vicinity of the grounding electrode on the disturbing circuit, however, the experimental observations of mutual impedance gradient can be explained fairly well in terms of a d.-c. distribution. The curves of Figs. 5 and 6 show that in the immediate neighborhood of the electrode, the gradient along any radius diverging from the electrode decreases very rapidly with increase in distance from the

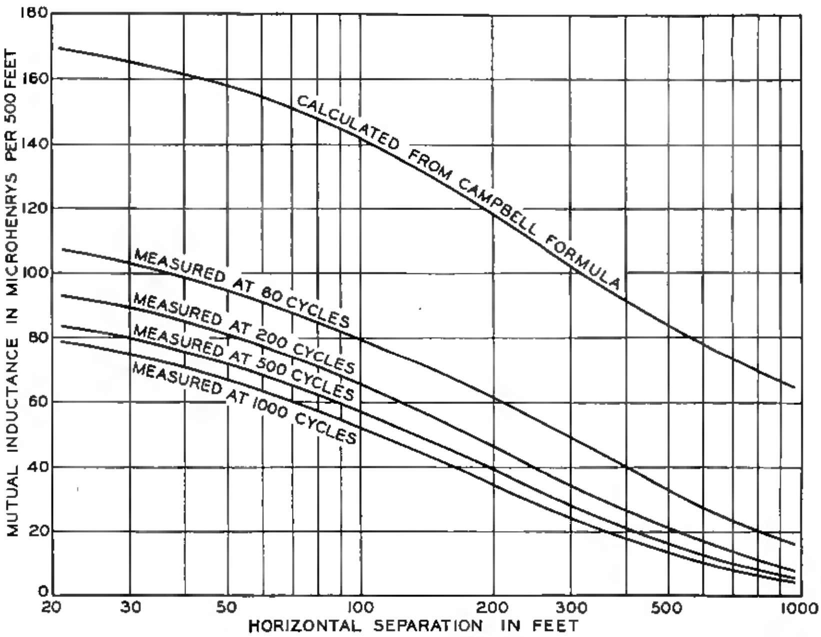


Fig. 9—Cross Keys tests. Campbell theory comparison of measured and calculated mutual inductances.

electrode, and is approximately in phase with the current. The gradient along the radius under the disturbing line approaches asymptotically a constant value, and beyond 300 ft. from the electrode the phase angle changes rapidly from a very small value to a value approximating 80 degrees. The gradient along the other radii, however, appears to decrease indefinitely and the phase angles are smaller. Such effects are in qualitative accord with predictions based on a d.-c. distribution, as will be seen by reference to Figs. 10 and 11.

On Fig. 10 are plotted the resistance and reactance components of the observed gradient under the disturbing line, with values computed using Campbell's formulas. Two calculated curves for the resistance component are plotted, for conductivities of 2.5×10^{-15} and 2.5×10^{-13} (abmhos per cm. cube). It will be seen that the experimental values lie between these two curves, tending towards the former for

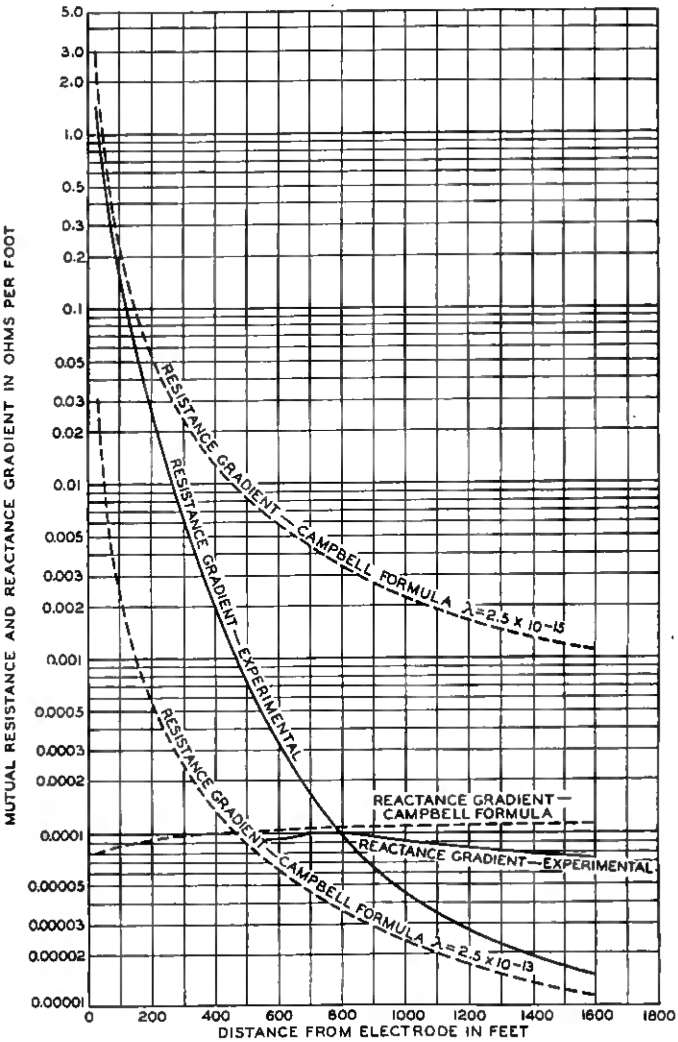


Fig. 10—Cross Keys tests. Mutual impedance gradient in vicinity of grounding electrode. Comparison with calculated values.

short distances from the electrode and towards the latter for long distances. As in the measurements previously described, the calculated mutual reactance component is greater than the measured value, although in this case the discrepancy is substantially smaller. On Fig. 11 are shown the phase angles of the gradient as computed from the calculated values of resistance and reactance components given on Fig. 10. Here also the measured curve falls between the two calculated curves.

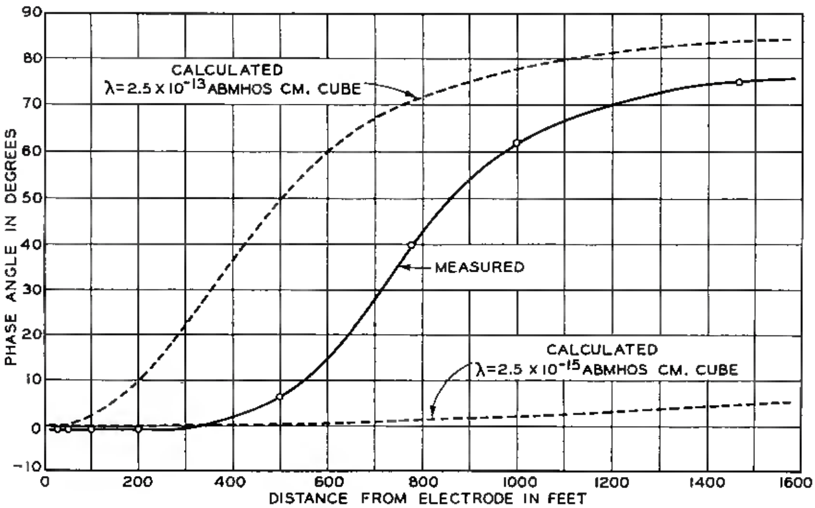


Fig. 11—Cross Keys tests. Comparison of measured and calculated (from Campbell's formula) angles of mutual impedance gradient.

Since the gradient near the electrode is obviously affected mainly by the conductivity of the earth in the immediate neighborhood, and that at remote points is influenced more by the conductivity of the earth at substantial depths, the possibility that the earth in this region is not homogeneous, but stratified, is suggested. These curves seem to support the conclusion that the earth in this neighborhood has at least two strata, the upper one having a very low conductivity and the lower one a conductivity approximately a hundred times greater. Further experimental evidence tending to the same conclusion has been obtained, and will be described presently. For the present it may be pointed out that this conclusion is supported by the geological data pertaining to this region, for which an upper layer of sand and gravel from 130 to 170 ft. in depth is indicated, superimposed on a mixed structure of sand, clay, and shale, with a substantial amount of ground water.

Methods Considering A.-C. Distribution of Earth Current. The problem of computing the mutual impedance of ground-return circuits, considering an a.-c. distribution in the earth has been attacked by several writers.[3] In the interpretation of the experimental results, the papers of J. R. Carson [7] and F. Pollaczek [6] have been used, since a minimum of assumptions was made in the solutions advanced by these writers. The assumptions made are that the disturbing circuit is straight and of great length,[9] that the propagation constant, in absolute units, of the circuit is very small compared to unity, and also that the earth is a homogeneous body of fairly good conductivity. With these assumptions it is found possible to solve the fundamental field equations for the magnetic and electric fields in the vicinity of the disturbing conductor at points remote from the ends of the circuit and thence to get the mutual impedance. Physically, this method recognizes and takes into account the fact that in a conductor of large extent, such as the earth, the distribution of alternating current will be influenced by the changing magnetic field. Qualitatively, the effects are similar to those involved in the well-known skin effect, and may be thought of in terms of a distribution of eddy currents in the earth. It is obvious that the distribution of the eddy currents will depend on the earth conductivity and also on the frequency. The resultant fields, and hence the mutual impedances, will then be functions of earth conductivity and of frequency.

Presentation of the formulas and graphs giving the results of the analysis is outside the scope of the present paper, and reference should be made to the original papers for these. As an illustration of the results, however, the curves of Fig. 12 have been prepared, showing the calculated mutual resistance and reactance of ground-return circuits at a frequency of 60 cycles for several values of earth conductivity, within the range of experimental values. Both the resistance and reactance components are seen to be pronouncedly affected by earth conductivity, particularly for the larger separations.

In applying this theory to the tests made at Cross Keys, the procedure adopted, in the absence of direct data on the earth conductivity at this location, was to choose an earth conductivity which would result in the best fit between the calculated and observed values, and to see whether a single value for earth conductivity would suffice to explain all the results. On Fig. 13 comparisons have been made between the experimentally determined mutual impedances for the 60- and 500-cycle frequencies; the curves were computed by use of the formulas given by Carson. It will be seen that in so far as the magnitude of the mutual impedance is concerned an excellent agreement can be made between

[3] See bibliography references 3 to 9, inc.
[7] See bibliography.
[6] See bibliography.
[9] See bibliography.

the calculated and observed values. However, for the best agreement it is found necessary to assume a different earth conductivity at 500 cycles than at 60 cycles. Thus, while at 60 cycles the indicated earth conductivity is 4.2×10^{-13} abmhos per cm. cube, at 500 cycles it is

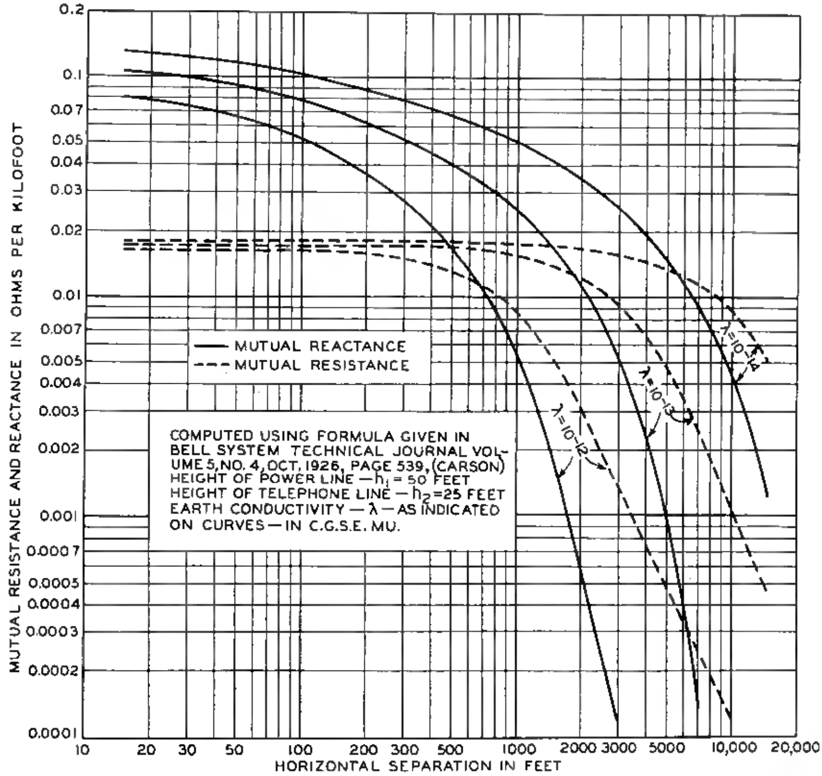


Fig. 12—Resistance and reactance of ground-return circuits. Frequency 60 cycles.

2.76×10^{-13}. However, a computed curve for 500 cycles, using a conductivity of 4.2×10^{-13} falls below the experimental curve by only 30 per cent. Table I gives the values of earth conductivity required to

TABLE I.

Cross Keys Tests.

Earth Conductivity Giving Best Agreement Between Calculated and Measured Values of Mutual Impedance.

Frequency cycles	Indicated earth conductivity from Carson's formulas abmhos per cm. cube
60	4.2×10^{-13}
200	3.75×10^{-13}
500	2.76×10^{-13}
1000	2.0×10^{-13}

give the best fit to the curve of mutual impedance at each frequency. The range is not large, extending only from 4.2×10^{-13} at 60 cycles to 2.0×10^{-13} at 1000 cycles.

Turning to the components of the mutual impedance, however, the agreement is found to be not as good. Fig. 14 shows the measured values of mutual resistance and reactance at 60 cycles and at 500 cycles, also the computed values, the calculations at each frequency being made with the earth conductivity indicated in Table I. At 60 cycles

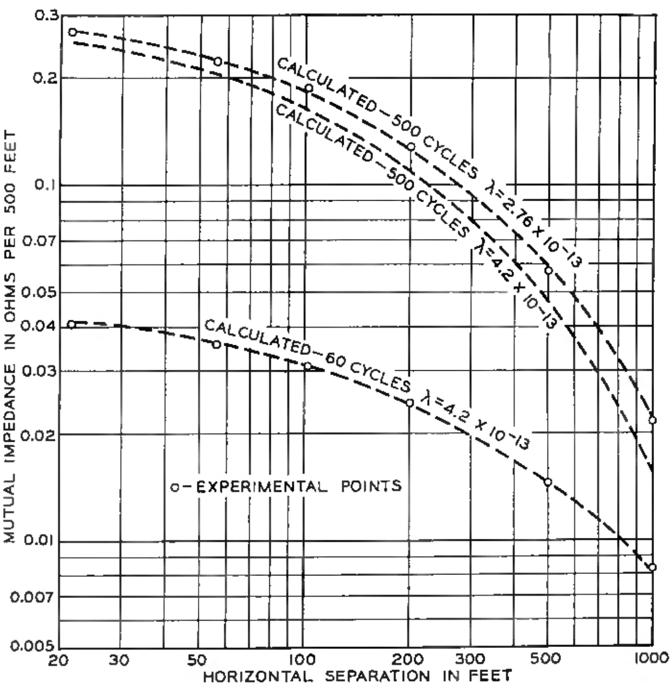


Fig. 13—Cross Keys tests—Carson theory. Comparison between calculated and measured values of mutual impedance.

the agreement is quite good, but at 500 cycles the departure between calculated and measured values is large. The measured mutual resistances are consistently lower than those calculated, while the measured mutual reactances are higher.

As indicated by the above comparisons, a theory of the kind under discussion leads to results which are in quite good quantitative agreement with the experimental results; it is of some interest to discover whether an extension of the theoretical ideas would lead to still closer agreement. It was stated previously that the measurements around the grounding electrode could be accounted for on the hypothesis that the earth in this neighborhood is stratified, with a conductivity of around 2.5×10^{-15} near the surface and 2.5×10^{-13} in the lower depths. Qualitatively, it is to be expected that with such an earth structure the mutual resistance would be less, and the mutual reactance greater, than the corresponding values for an earth of uniform conductivity, since the eddy currents near the surface of the earth will be less, due to the lower earth conductivity.

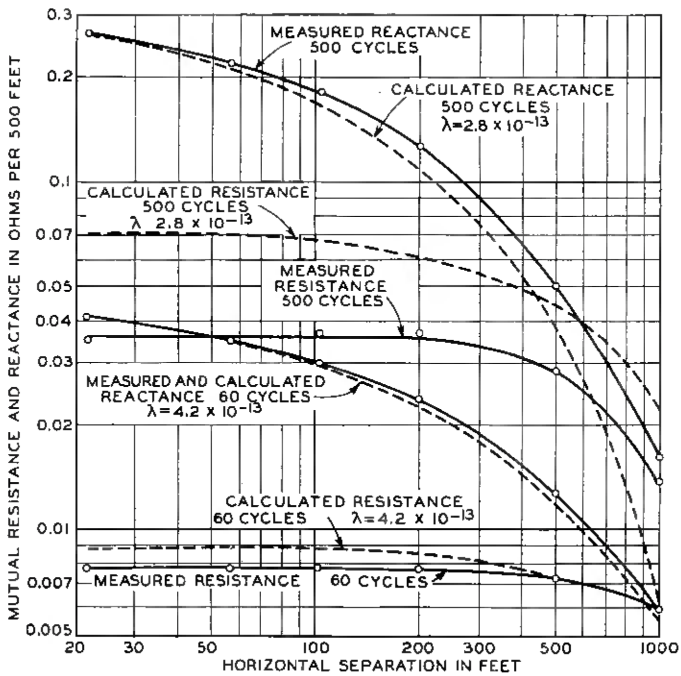


Fig. 14—Cross Keys tests—Carson theory. Comparison between calculated and measured values of mutual resistance and reactance.

Quantitatively, it would seem that a first approximation to the effect of a stratified earth in which the upper stratum has a much lower conductivity than that of the lower region could be obtained by assuming that the currents in the upper layer are negligible and hence that this layer can be abolished. The mutual impedances can then be worked out by the formulas applicable to a homogeneous earth, using the earth conductivity of the lower region and fictitious conductor heights, formed by adding the thickness of the upper stratum to the heights of the conductors above the actual earth's surfaces. Preliminary calculations have been made using this scheme, and it was found that using a conductivity of 5.0×10^{-13} for the lower stratum and a thickness of 130 feet for the upper stratum an excellent agreement could be found between calculated and observed values for all frequencies. The agreement extended not only to the magnitudes of the mutual impedances, but to the components as well.

Because of the simplifying assumption that the disturbing circuit is so long that the effects due to the ground connections at its ends can be neglected, the theory which we have been discussing is obviously inadequate to explain the phenomena in the neighborhood of the grounding electrode.

Practical Methods for Predetermining Coupling between Power and Telephone Lines

The ultimate purpose of the work of the committee is to develop simple methods to enable the calculation of the mutual impedance between power and telephone circuits before they are built. It is evident from the foregoing discussion that the use of any formula for the mutual impedance of ground-return circuits requires a knowledge of the conductivity of the earth or of the depth of the equivalent ground-plane. In the relatively few places in which tests have been made, a range of earth conductivity from 10^{-12} to 10^{-14} abmhos per cm. cube has been observed, and reference to Fig. 12 indicates that within this range of earth conductivity a variation in mutual impedances of 20 to 1 or more may exist. Therefore, other experimental work has been done with the object of developing relatively simple testing schemes, the results

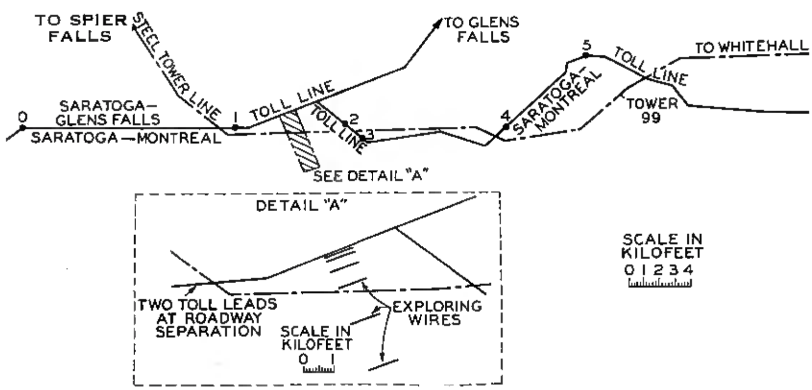


Fig. 15—Glens Falls tests.

	Circuit arrangements	
Tower 99	3-Phase conductors	Grounded
Circuit	Power	Telephone
Height	50 ft.	20 ft.
Frequency	60 cycles	

of which could be used to predict the coupling coefficients in advance of the construction of the power or telephone line. An obvious method is to determine an experimental coupling curve by performing tests similar to those made at Cross Keys, using short-length disturbed circuits and either an existing power or telephone line, or a specially laid out conductor, as the disturbing line.

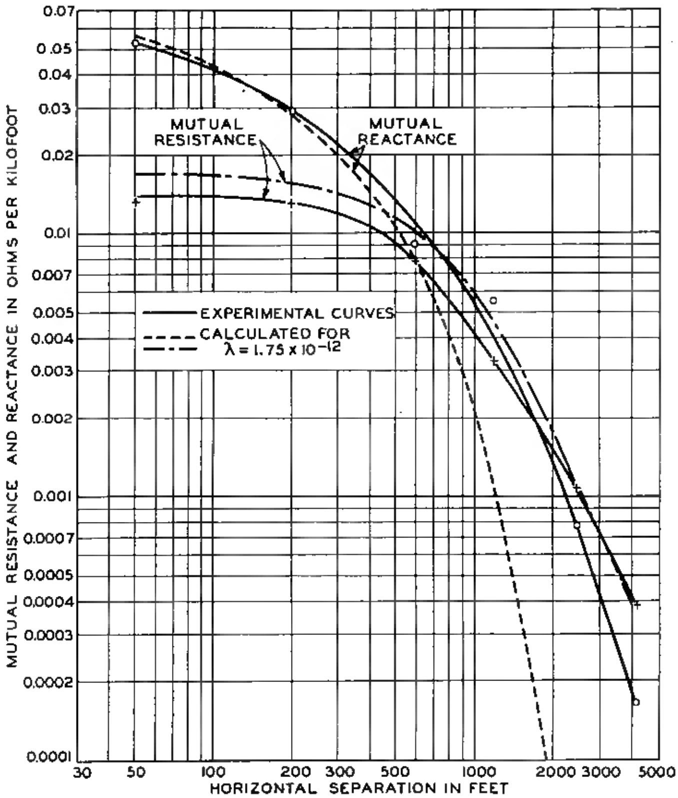


Fig. 16—Glens Falls tests. Experimental values of mutual resistance and reactance.

This experimental curve would then be used to compute the coupling between power and telephone lines. One advantage of using an experimentally determined coupling curve is that it obviates the necessity of knowing or assuming a structure and conductivity of the earth; the coupling curve can be used directly without reference to any theoretical formulas. To determine the practicability of such a scheme, 60-cycle tests have been made in two locations where existing exposures were present, for the purpose of determining the accuracy with which experimental observations could be predicted.

Tests at Glens Falls, N. Y. Fig. 15 shows the arrangement of circuits involved in tests made at Glens Falls, N. Y. A section of the Saratoga-Glens Falls telephone line about six miles in length was energized with ground-return current. Measurements were made of the voltages induced in short ground-return circuits laid on the ground parallel to the straight section of the telephone line. The resistance and reactance components of the mutual impedance derived from these measurements are given on Fig. 16. As a matter of interest the mutual resistance and reactance computed by the use of the Carson formulas for an earth conductivity of 1.75×10^{-12} are also given. This earth conductivity gives the best agreement between the calculated and observed magnitudes of the mutual impedances. The general agreement between the computed and observed quantities is much like that found from the Cross Keys tests.

Earth return current was then sent over the power line from Spier Falls to Tower 99, and induced voltages measured in the entire exposed section of the Saratoga-Montreal telephone line, and in several parts of the exposure as indicated on the sketch. In Table II, the observed

TABLE II.

GLENS FALLS TESTS

Measured Mutual Impedances of Power and Telephone Circuits and Comparison with Values Calculated from Coupling Curves of Fig. 16.

Section of telephone line	Measured mutual impedance—ohms	Calculated mutual impedance—ohms
0–1	.0586 ∠68.5°	.0614 ∠53.3°
1–2	.0294 ∠52.4°	.0564 ∠49.8°
2–3	.0476 ∠73.4°	.0382 ∠69.6°
3–4	.107 ∠56.4°	.113 ∠49.2°
4–5	.100 ∠44.4°	.0117 ∠35.8°
0–5	.347 ∠5[illegible].8°	.267 ∠55.3°

mutual impedances determined from this latter test are compared with values calculated by using the experimental coupling curve given on Fig. 16. The agreement between computed and observed values is, in general, only fair, although for two of the parts, the agreement is excellent. It is thought that the rather poor check for Sections 1-2 and 2-3 is due to the inductive effect of currents set up in the ground wire on another power line which extended through these sections. With regard to the extreme departure of the measured mutual impedance for Section 4-5 from that calculated it is impossible to decide the cause from the experimental data available. A possible explanation is that it is due to a large difference in earth conductivity in this region from that in the region in which the coupling curve was determined. The large difference in this section is reflected in the rather poor check in the overall coupling (Section 0-5).

Tests at Massillon, Ohio. Tests made at Massillon, Ohio, were similar to those at Glens Falls, except that the arrangements were somewhat more elaborate. The layout of the circuits involved is shown in Fig. 17. The exposure is about 16 mi. long with separation between the power and telephone line ranging from a crossing to about 4200 ft., a large part of the exposure being at a separation of about 3000 ft. A set of "exploring wires," each 200 ft. in length, was laid on the ground parallel to the telephone line as shown in the detail of Fig. 17. These were arranged in four groups and were distributed over an area approx-

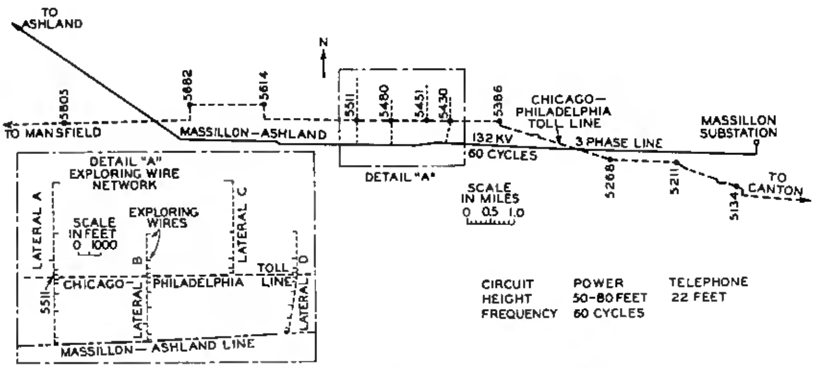


Fig. 17—Massillon tests—circuit arrangements.

imately 1½ by 2 mi. Coupling curves were determined from measurements of the voltage induced in the exploring wires for the condition of the telephone line energized with 6 amperes ground-return current, and also for the condition of the power line energized with 40 amperes ground-return current. The mutual impedances derived from the two sets of measurements are practically identical. Fig. 18 shows the resistive and reactive components of the coupling curve using the average of all measurements made on the exploring wires. A comparison of the measured curves with curves calculated by Carson's formulas for a value of earth conductivity of 3.6×10^{-13} abmhos per cm. cube show the same type of agreement as that observed at Cross Keys and Glens Falls.

The principal reason for using such a large number of exploring wires on this particular test was to investigate the effect of local irregularities of the earth upon an experimental coupling curve and to determine the minimum number and length of exploring wires which it is necessary to use in order to be reasonably confident of the accuracy of the results.

The data indicate that if only one of the seven groups of 200-ft. exploring wires had been used, the maximum deviation of any one point from the corresponding point on the average curve would have been less than 25 per cent and that the probable deviation would have been less than 10 per cent. This deviation could probably be reduced by using a somewhat longer exploring wire.

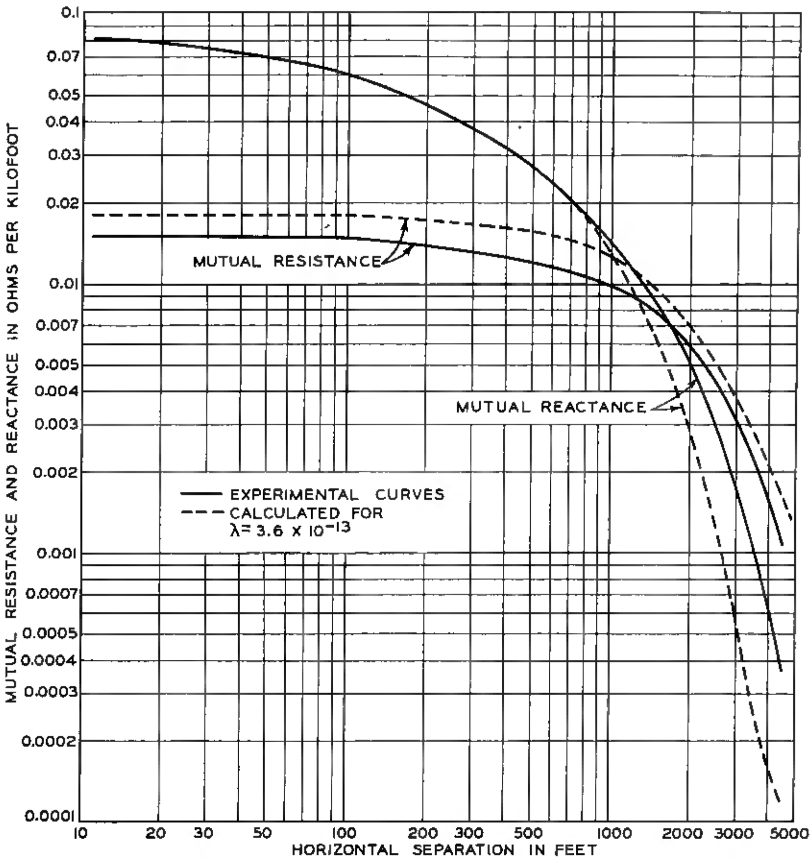


Fig. 18—Massillon tests. Experimental values of mutual resistance and reactance.

Measurements were also made of the voltage induced in sections of the telephone line when the power line was energized with ground-return current. A comparison of the measured values of the mutual impedance and those given from calculations using the coupling curves of Fig. 18 are given in Table III. With the exception of the section from pole 5134 to 5211, for which the calculations may be subject to some error due to proximity to the end of the disturbing circuit, and the section from pole 5614 to 5682, the agreement between measured and calculated values is statisfactory.

TABLE III.

MASSILLON TESTS.

Measured Mutual Impedances of Power and Telephone Circuits and Comparison with Values Calculated from Coupling Curves of Fig. 18.

Exposure		Measured mutual impedance-ohms	Calculated mutual impedance—ohms
From pole	To pole		
5134	5211	$0.0308 \angle 45.9°$	$0.0415 \angle 35.5°$
5211	5268	$0.0532 \angle 44.7°$	$0.0442 \angle 34.2°$
5268	5386	$0.282 \angle 63.7°$	$0.280 \angle 64°$
5386	5511	$0.0707 \angle 30.3°$	$0.0630 \angle 28.2°$
5511	5614	$0.0386 \angle 29.7°$	$0.0427 \angle 28.5°$
5614	5682	$0.0164 \angle 18.4°$	$0.0117 \angle 14.8°$
5682	5805	$0.141 \angle 62.4°$	$0.149 \angle 59.2°$
5134	5805	$0.609 \angle 53.7°$	$0.612 \angle 52.5°$

D.-C. Determination of Earth Conductivity. In considering the experimental results described above, particularly the reasonably good agreement between the experimental coupling curves and those calculated by means of the Carson formulas with suitably chosen earth conductivity, it seemed desirable to investigate whether an experimental value of earth conductivity alone would be sufficient information for the determination of coupling curves with enough accuracy for many purposes. With this in mind, an investigation has been undertaken of more direct methods for determining the conductivity of the earth or more generally of methods for determining the structure of the earth in a particular location (whether homogeneous or stratified and, if the latter, the thickness of the strata) and the earth conductivity. The work is at present only in an early stage, but a brief statement of the method followed and of the results so far obtained may be of interest.

The procedure followed amounts to an investigation of the mutual resistance of a number of suitably located ground-return circuits, with direct current. It will be recalled that in an earlier part of the paper it was stated that for a homogeneous earth the mutual resistance of two ground-return circuits, for direct current, can be easily derived, and expressed in a formula involving only the earth conductivity and the distances between the grounding electrodes. For a stratified earth, similar formulas have been worked out involving the distances between the grounding electrodes and the conductivities, and the thicknesses of the several strata. By means of measurements of the mutual resistance for direct currents of circuits with suitably located ground electrodes, the conductivities and thicknesses of the strata can then be determined.

Practically, the experimental work presents many problems, among them being the elimination of the effects of stray earth currents and evaluation of the effects local irregularities in earth conductivity. A preliminary trial of the method was made in connection with the tests at Massillon, and while local irregularities were found to be quite marked, yet the average earth conductivity in the region covered by the tests was about 1.5×10^{-13} abmhos per cm. cube, which is not greatly different from that indicated by the coupling tests. A quite extended series of tests at Cross Keys, using an improved technique, yielded results in excellent agreement with the hypothesis that at this location the earth is stratified, having an upper layer about 150 ft. thick with a conductivity about 3.4×10^{-15} and a conductivity in the lower stratum about 2.6×10^{-13}.

Conclusions

In conclusion, it is well to recall the end towards which the work described in this paper has been directed. It was desired, first, to obtain a sufficiently detailed experimental study of the mutual impedances of ground-return circuits to enable the formation of an adequate picture of the physical phenomena involved; also to test out the theoretical formulas available. Second, the aim was to investigate practical means for enabling the calculation of the ground-return mutual impedances of power and telephone lines.

With regard to the first item, it was found that an analysis in terms of an "equivalent ground-plane" was inadequate to represent completely the observed phenomena. However, when information is available as to the proper value of ground-plane depth, this method can be used to advantage in many cases where approximate results only are desired.

A theory based on the assumption of a d.-c. distribution in the earth gave a somewhat better explanation, particularly in connection with the mutual impedances of circuits in which the points of ground connection were in close proximity, but left much to be desired in the way of quantitative agreement with the experimental results. The results of a theory which considers the effect of eddy currents in the earth are shown to be in fair qualitative agreement with some of the test values, and by a slight extension of the theory, good quantitative agreement can be found. This theory, however, does not explain end effects.

In the investigation of practical means for enabling the calculation of the mutual impedances of power and telephone lines a scheme involving the experimental determination of a coupling curve has been found to give quite satisfactory results. Further work is to be done on this problem, and similar tests must be made in several other locations before the method can be considered completely satisfactory. Other methods are also being investigated.

Acknowledgment

The authors wish to express their appreciation of the assistance of Messrs. F. J. Grueter and B. C. Griffith in the field testing, analysis of the data and preparation of this paper, also to many others who have aided in this work. Thanks are due to the operating power and telephone companies who lent the facilities for some of the work described, and to members of their organizations who assisted in the tests.

Bibliography

The following list is far from complete, but covers the more important papers which were consulted in the course of the work:

1. California Railroad Commission, Inductive Interference, 1919.
2. G. A. Campbell, "Mutual Impedances of Grounded Circuits," *The Bell System Technical Journal*, 2, October 1923, pp. 1–30.
3. F. Breisig, "Ueber die Berechnung der Magnetischen Inducktion und Wechselstromleitungen mit Erdruckleitung," *Telegraphen und Fernsprechtechnik*, 14, April 1925, pp. 93–98.
4. O. Mayr, "Die Erde als Wechselstromleiter," *E. T. Z.*, 46., September 1925 pp. 1352–1436.
5. R. Rudenberg, "Die Ausbreitung der Erdstrome in der Umgebung von Wechelstromleitungen," *Zeit. fur Angew. Math. und Mech.*, 5, October 1925, pp. 361–389.
6. F. Pollaczek, "Ueber das Feld einer Unendlich Langen Wechelstrom Durch Flossenen Einfachleitung." *Elektrische Nachrichten Technik*, 3, September 1926, pp. 339–359; 4, January 1927, pp. 18–30.
7. J. R. Carson, "Wave Propagation in Overhead Wires with Ground Return," *The Bell System Technical Journal*, 5, October 1926, pp, 539–554.
8. G. Haberland, "Theorie der Leitung von Wechelstrom Durch die Erde," *Zeit. fur Angew. Math. und Mech.*, 6, October 1926, pp. 366–379.
9. F. Pollaczek, "Gegenseitige Induktion Zwischen Wechselstromleitungen von Endliche Lange, *Ann. der Physik*, Vol. 87, Dec. 1928, pp. 965–999. This paper treats of the mutual impedances of circuits of finite length. However, the assumptions regarding conditions at the ground connection seem to depart considerably from the practical conditions of the experiments reported in the present paper. For this reason, and the fact that numerical evaluations by these formulas are very laborious, no calculations by this method have as yet been made for comparison with the experimental results.

A Survey of Room Noise in Telephone Locations *

By W. J. WILLIAMS and RALPH G. McCURDY

This paper describes a survey made to determine the range of magnitudes of room noise present in telephone locations. Measurements were made in a total of 250 locations in New York City and environs, distributed among businesses and residences in accordance with telephone traffic distribution. In each location, measurements were made by a marginal audibility method using the human ear as a part of the measuring device, and by a visual indicating meter. A brief description is included of the apparatus employed with each of these methods. Results are presented for measurements made in various classes of rooms, under winter and summer conditions.

AMONG the projects of the Joint Subcommittee on Development and Research of the National Electric Light Association and Bell System is one (No. 4) which is studying the effects of noise [1] on telephone transmission and methods for its measurement. It was appreciated that, in addition to noises of electrical origin caused by exposures to power circuits or by sources incidental to the operation of the telephone system, there are also noises in the rooms in which telephones are used which have an important effect on telephone service. In studying the effects of noises, it is, of course, necessary to consider both noises of electrical origin and room noises. It was desired that, in laboratory tests of the effects of line noises on speech transmission, typical amounts of room noise should be provided at the test location. The survey described herein was made to obtain room noise data for these laboratory tests.

The methods described should be of general interest in connection with other noise problems. Increasing attention is being given, both in America and in Europe, to the general problem of noise as an undesirable attribute of modern civilization. Some efforts are being made to investigate sources of city noise. Modifications have been made in the design of machines and appliances, such as typewriters, motor cars, electric refrigerators, rotating electrical machinery, and domestic oil burners, so as to reduce the noise involved in their operation. Attention is being given to the quieting of rooms by means of acoustic treat-

*Presented at the Summer Convention of the A. I. E. E., Toronto, Ont., Canada, June 23–27, 1930.

[1] In this joint work, noise is taken to mean any extraneous sound which would tend to interfere with telephone conversation.

Room noise is used to include any extraneous sounds at the place where the measurement is made, except those proceeding from the telephone receiver. It thus includes, in addition to noises such as the rattling of papers or the roar of street traffic, any other sounds extraneous to the telephone conversation, for example, those of other conversations or of music produced nearby.

ment. Studies are being made of the effects of noise on living beings, including effects on the efficiency of workers.[2] In all of this work, quantitative measurement is important.

For the specific problem in hand, it was desired to obtain information on the magnitudes of room noises, as well as some general indication of the frequency composition of typical room noises.

While it is recognized that ordinary room noise is a highly variable quantity changing from instant to instant in intensity and frequency composition, it is felt that sufficient measurements were made to specify the makeup of a typical room noise for use in the laboratory tests, and in addition to obtain an indication of the effect of various factors, described below, upon the noise. Since it was desired to make the measurements as representative as possible of typical telephoning conditions they were made at times of day and in types of locations determined by a study of telephone message traffic. Since the results would be affected by the choice of locations, they are presumably less typical for non-telephone than for telephone purposes.

The residences included in the survey ranged from apartments in large city buildings to small homes in outlying towns. In the business locations were included offices, stores, factories and workshops, and public buildings, such as hotels and clubs.[3] Establishments of various characters were included in each classification; they ranged in size from small stores to great manufacturing plants.

In making all measurements, an attempt was made to simulate the normal conditions which would obtain when a telephone call was placed. If noises existed in the room, which would be discontinued when the telephone was being used, such noises were stopped while the measurements were being made. On the other hand, care was taken to see that none of the normal noises of a particular location was discontinued because of the fact that measurements were being taken.

It was recognized that there would be a difference between the room noise experienced on local and on long-distance calls. The survey was made on the basis of telephone traffic as a whole, which consists predominantly of local calls.

The survey consisted of two series of tests, one made during the months of January, February, March and April, and the other made during the months of July and August. The former series was the more comprehensive, including 205 measurements; the results given herein are based on this series of tests except where specifically noted

[2] D. A. Laird, "The Effects of Noise," *Jl. Acoustical Soc. Amer.*, Jan. 1930, p. 256.

[3] In public buildings, only a very small proportion of the telephone locations tested were in booths or at coin-box telephones.

otherwise. The second series of tests was made for the purpose of determining the difference between the room noise encountered under winter conditions and that encountered under summer conditions. Consequently, a selected group of the locations, which had been measured in the winter, were measured again under summer conditions.

It must be appreciated, in generalizing from the data given, that tests were made in only a limited number of locations.

Two methods were employed in making the measurements described in this paper, one electrical and the other aural. The electrical method employed a condenser-transmitter pick-up, amplifiers, and detector. A weighting network was incorporated in the amplifier to simulate the sensitivity characteristic of the ear. The aural method, known as the "masking method," involved the measurement of the masking effect of the noise on various warbler tones recorded on a phonograph record. Both of these methods will be described in greater detail below.

General Results

Some of the interesting results which were obtained from this survey may be summarized as follows:

On the average, room noise in residences was about 20 db less in magnitude than that in business locations.

The spread in the magnitudes of business room noises was about 40 db, as compared to 20 db for residence room noises. These spreads include 90 per cent of the measurements, excluding the lowest and highest 5 per cent. The standard deviation of the measurements was about 12 db for business noise and 6 db for residence noise.

Room noises averaged 4 or 5 db higher in summer than in winter.

In general, the magnitude of residence noise was affected to only a minor extent by the size of the town or city in which it was measured.

On the average, the frequency composition of residence noise was about the same as that of business noise. The masking effect of the noise on a tone covering the range 750-1500 cycles was greater than that on ranges above and below this. The magnitudes of components in the lower part of the range covered (about 250-5000 cycles) appeared to be somewhat larger than those in the higher part of this range.

Methods of Measurement

The two methods which were employed in the survey are as follows:

Aural Method—Masking of Warbler Tone.[4]—In this method a tone of varying pitch (warble) is produced and sent into a receiver. The receiver cap is provided with slots shaped so that the observer's ear

[4] R. H. Galt, *Jl. Acoustical Soc. Amer.*, October 1929, p. 147.

canal is always open to the air of the room regardless of how firmly the receiver is pressed against the ear. The tone is generated by means of a phonograph record and a magnetic phonograph record pick-up, and is a variable-frequency tone, the pitch of which varies between certain limits several times per second. An attenuator is placed between the magnetic pick-up and the receiver. The observer sets the attenuator at a point where he can barely recognize the sound of the warble in the presence of the room noise. He also obtains the setting at which he can barely hear the warble in a perfectly quiet room. The difference between these two settings is a measure of the masking effect of the noise in this room upon the warbler tone, for this particular observer.

An idea of the frequency composition of a given room noise may be obtained by using several different warbler tones, each covering a different portion of the voice-frequency range. This is based on the fact that, in general, a tone of a given frequency masks to a greater extent tones that are near it in frequency than tones that are far removed from it in frequency.

The phonograph records used in the present room noise survey were three-band records, i.e., three warbler tones were cut on each record, each tone occupying about one-third of the available space. The frequencies included in the various bands were as follows: high band, 1500–5600 cycles per second; middle band, 750–1500 cycles per second; low band, 250–750 cycles per second. In each band the frequency varied continuously from the lower to the upper limit and back to the lower limit, the period of such a complete "warble" being about one-sixth of a second.

Electrical Method—Room Noise Meter.—There is, of course, a number of different electrical methods which might be employed for measuring room noise, ranging from a single over-all measurement to a complete wave shape or frequency analysis. The complete analysis or the measurement of energy present in a considerable number of narrow frequency bands is subject to the disadvantages, for such a survey as this, of slowness of measurement and bulkiness of testing equipment.

The method which was adopted was one based on the use of a frequency weighting simulating the sensitivity of the ear. This frequency characteristic is shown on Fig. 1. It is an equal loudness weighting; that is, the room noise meter was so constructed that different single-frequency noises of equal loudness would give approximately the same meter readings. The shape of an equal-loudness curve is somewhat flatter for high levels of loudness than for low ones. The weighting curve chosen for the meter was for a loudness corresponding to that of a 1000-cycle tone 30 or 40 db above the threshold of audibility. This general level is not far from the middle of the range of levels of room noise components. The loudness data used were those given by Kingsbury,[5] based on experimental data on single-frequency tones.

The sensitivity of the meter is such that a 1000-cycle tone about 28 db above threshold would give a reading of 0 db on the meter scale.

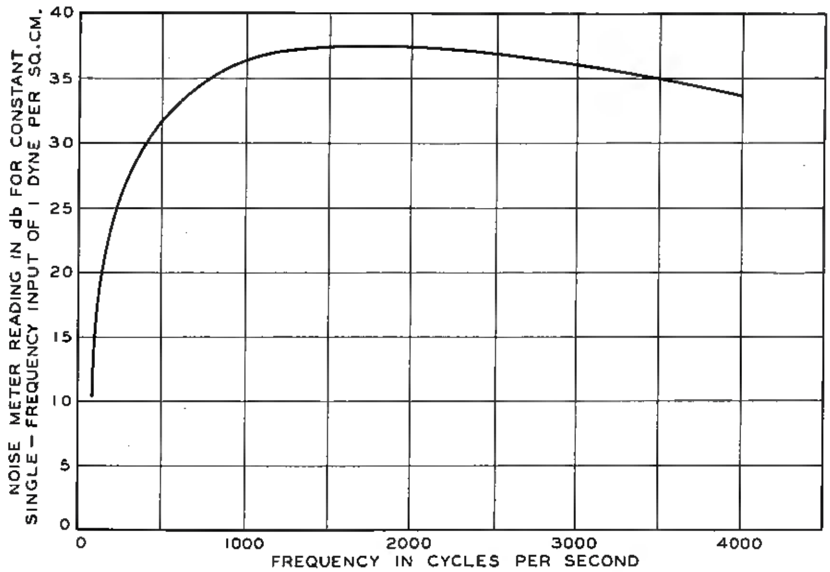


Fig. 1—Response characteristic of room noise meter.

The room noise meter employed is shown together with its auxiliary equipment in Fig. 3. It consists of a condenser transmitter for converting acoustical energy into electrical energy, six stages of amplification for raising the level of the noise currents sufficiently to operate a thermocouple meter indicating device, and a weighting network, as de-

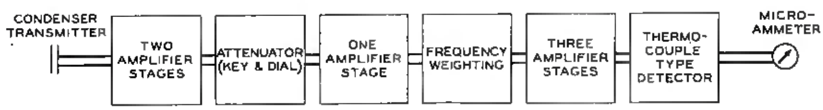


Fig. 2—Schematic diagram of room noise meter.

scribed above, as well as certain apparatus not employed in obtaining the results reported here. The general layout of the circuit is indicated in the schematic diagram of Fig. 2. A portable battery supply and means for calibrating form the necessary auxiliary equipment. An

[5] *Physical Review*, Vol. 19, April 1927, pp. 588–600.

adjustable attenuator controlled by a key and a dial is provided between stages of the amplifier so that the noise energy being measured may be brought within the range of the meter over a range of levels of 80 db (corresponding to a power range of 100,000,000 to 1).

Fig. 3—Room noise meter and auxiliary equipment.

Operation of the Room Noise Meter.—The noise meter is first calibrated, as described below, so that its sensitivity is set at a predetermined value. The condenser transmitter is then placed at the spot where it is desired to measure noise, and the gain of the amplifier is adjusted by means of a key and dial until the needle of the microammeter in the output circuit fluctuates about a given point. The settings of the key and dial then give a measure of the noise. In addition to the average readings obtained in this manner, readings of the fluctuations in the noise can be similarly obtained. As an aid in the reading, the microammeter scale is calibrated in decibels.

The calibration of the meter in the field consists of a check on the over-all sensitivity of the instrument. The filament currents and plate voltages are adjusted to the correct values. Then a fixed percentage of the electrical output of a standard buzzer, the current from which is measured by a thermocouple, is fed into a special receiver which is placed in a prescribed way on the condenser transmitter. The gain of the amplifier is then adjusted until the output microammeter needle reaches a predetermined point. The sensitivity of the meter will then be as shown on Fig. 1.

An over-all calibration of the meter, as a function of frequency, is given on this figure. To obtain this, separate determinations were made of the volts generated by the condenser transmitter per unit of pressure, and the meter reading per volt generated by the transmitter, as a function of frequency; and the results were combined to give the values shown. Harmonics in the testing waves were reduced to such a point that they did not affect the results. After a substantial part of the survey had been completed, a check was made of the electrical portion of the calibration, and the changes found were quite negligible.

Accuracy of the Meter.—The precision of the apparatus is substantially greater than the precision with which ordinary varying noises can be measured. The readings obtained for steady inputs are proportional to the input, with an error of less than 1/2 db, over the entire range of noise amplitudes found in the survey. The apparatus is shielded electrically. In only one case did electrical fields produce any observed errors in the readings; this was when an attempt was made to measure the room noise near a rotary converter in a power station. The vacuum tubes are mounted in such a way that the effects of ordinary mechanical vibration on the readings are negligible.

Comparison of the Two Methods.—In general, the meter method gives results in physical terms while the masking method gives them in terms of effects on the ear; consequently, the choice of the method to be employed in any particular case depends somewhat on the use to which data will be put. It is true that the meter includes a network to simulate the sensitivity of the ear for various frequencies; it does not, however, simulate other properties of the ear, such as the departures from linearity in response by which subjective tones are produced by the ear mechanism, and the complicated way in which one sound masks another.[6]

The meter method, unlike the masking method, avoids any errors due to variations in human ears. This advantage is offset to some extent by the fluctuations of the meter needle, which make it difficult to obtain the mean reading if the noise is unsteady as is the case with most room noises.

In the case of noises of a distinctly intermittent, staccato character, the warbler tone can be heard and recognized in the brief intervals when the noise is a minimum. A preliminary investigation showed that, for a noise of this sort, the relation between readings obtained by the masking method and by the meter method was different from the relation obtained for a steady noise, the warbler readings being relatively lower in the case of intermittent noise.

[6] R. L. Wegel and C. E. Lane, "Auditory Masking and Dynamics of the Inner Ear," *Physical Rev.*, Feb. 1924.

Both methods were used in the survey, because it was felt that each gave information which could not be as accurately obtained from the other, and also because the use of two methods enabled each one to be used as a check upon apparatus defects which might occur in the other.

In using the masking method, data were taken by two experienced observers and corresponding measurements averaged. All meter measurements were made by one observer.

Results of Survey

Noise in Business Locations.—One hundred and nine business locations were visited. The magnitudes of the noises measured varied from that found in a doctor's quiet office to the din of a large manufacturing plant. Distribution curves for the noises measured are shown in Fig. 4 for the meter method and Fig. 5 for the masking method. For any point on one of these curves the corresponding per cent of all of the measurements made had values equal to or greater than the indicated abscissa value.

It may be seen that with the exception of the "high" curve of Fig. 5 the curves for meter and masking methods are fairly similar in shape. The "middle" curve has been selected to represent the masking method.

If there are excluded as extremes those noises which were so low that 95 per cent of all the noises measured equaled or exceeded them, and those which were so high that only 5 per cent of the measurements equaled or exceeded them, the spread of noise magnitudes is seen to be about 40 db. The standard deviation of the measurements is about 12 db.

As shown on Figs. 4 and 5 the median business room noise would produce a reading of 23 db on the meter scale and a masking of 27 db on the high-frequency warbler tone, 39 db on the middle-frequency tone, and 31 db on the low-frequency tone. The average business room noise was about 2 db higher than the median.

Some conception of the amounts of noise represented by these figures may perhaps be gained from the following. The extremely loud noise measured in a local station of the New York subway while an express train was passing produced a meter reading of 70 db, while the lowest noises measured in the survey, in quiet residences, gave readings near 0 db.

Data on noise at the business locations tested have been grouped so as to show the average differences in the room noise values obtained for different types of business and for different sizes of towns. It will be appreciated that only a very small number of measurements were in-

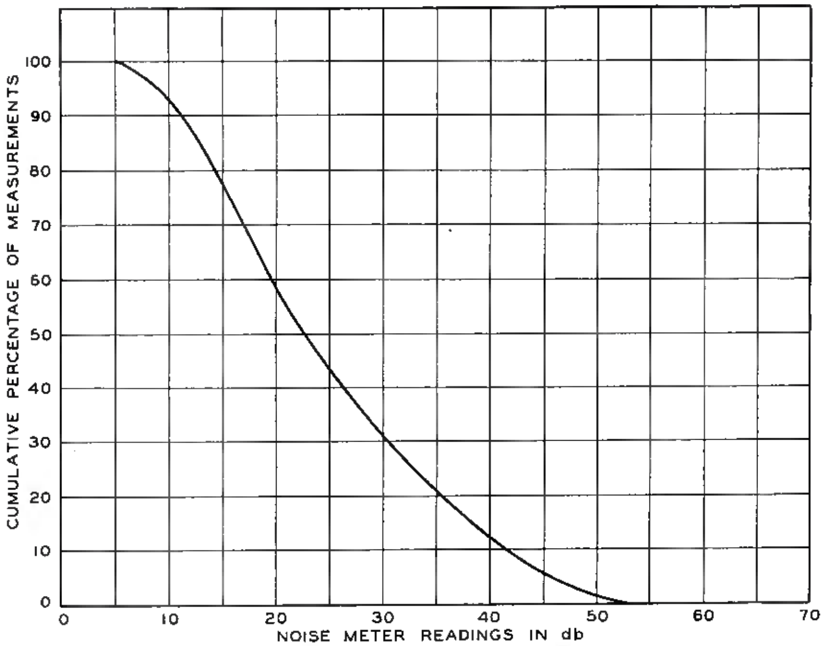


Fig. 4—Results of noise meter measurements of noise in business locations.

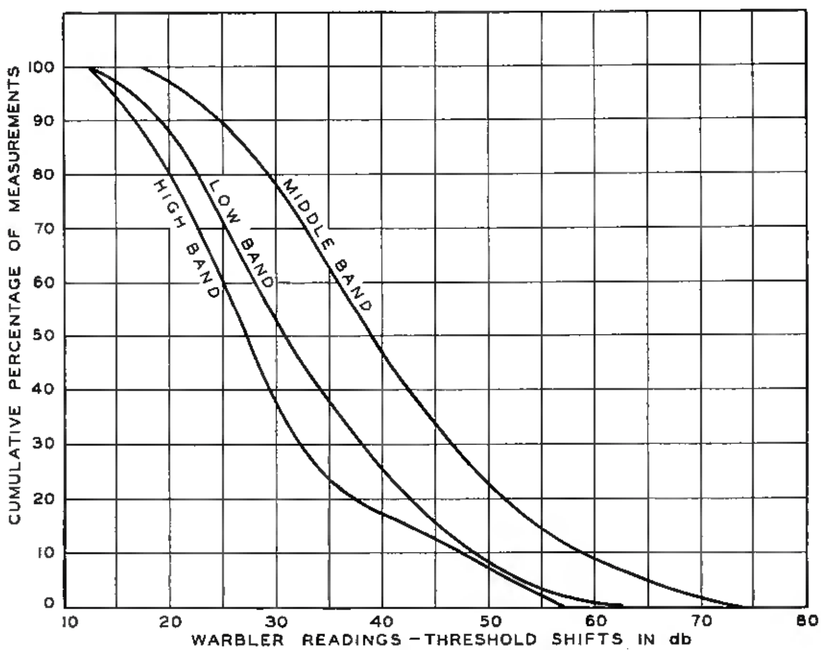


Fig. 5—Results of measurements of noise in business locations by masking method.

cluded in each sub-classification, and that consequently it is not safe to generalize from these sub-groupings as to room noise conditions in general.

Averages of the room noise measurements for the different types of business locations are shown in the following table.[7]

Type of business location	Masking of middle-frequency tone	Meter reading	Number of measurements
Offices	42 db	24 db	34
Stores	34	18	34
Factories	57	40	18
Public buildings	35	21	23
Average of all businesses (weighted according to number of measurements made)	40	25	109

The above figures show a significant difference between the noise measured in factories and that measured in other types of location. The other differences shown were found not to be significant when examined in the light of the spread in values for individual locations in each class.

Averages of the business room noise measurements obtained in various sizes of towns are shown in the table below.

Size of town	Masking of middle-frequency tone	Meter reading	Number of measurements
Class A (over 400,000 pop.)	45 db	26 db	39
Class B (100,000 to 400,000 pop.)	37 "	22 "	18
Class C (10,000 to 100,000 pop.)	42 "	27 "	41
Class D (less than 10,000 pop.)	27 "	11 "	11

These figures indicate that (with the exception of Class C towns) the business noise measured in large cities was greater than that in smaller towns. This is believed to hold true despite a fairly large spread in individual measurements within a given class. The exception in the case of Class C towns is explained by the fact that a fairly large percentage of the measurements in this class were made in large factories.

Room noise in business locations was observed to be quite complex in

[7] It will be noted that in the results given, the difference between the masking of the middle-frequency tone and the meter reading is relatively constant. It was found that for any considerable sub-group of the measurements, this difference was not far from 15 db. This figure, of course, would not in general hold for a single noise selected at random. There was a general tendency for the difference to be somewhat larger for larger values of noise.

frequency composition. The masking effect of the noise on the middle band was greater than that on the high and low bands. In order to give an approximate interpretation of this in terms of pressures in various frequency regions, account must be taken of the relative magnitudes of threshold pressures in the three warbler frequency bands, since the masking effects were obtained by subtracting threshold settings of the attenuator from the settings made in the presence of the noise. For

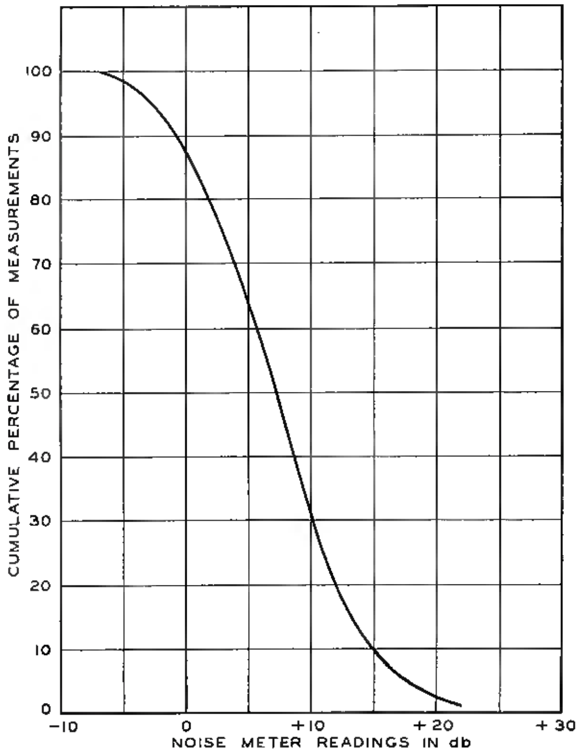


Fig. 6—Results of noise meter measurements of noise in residence locations.

the middle and upper bands, threshold pressures are about the same; hence, the lower values of masking for the high range indicate that components in this range are in general relatively weak. As previously determined,[8] threshold pressures at frequencies in the low band are several decibels higher than those in the other bands. Combining the values of masking for the low and middle bands with the corresponding threshold pressures, it is seen that the physical magnitudes of compo-

[8] H. Fletcher, "Useful Numerical Constants of Speech and Hearing," *Bell System Tech. Jl.*, July 1925.

nents in the low- and middle-frequency ranges are in general not far different. The above analysis is, of course, very rough, as the whole range from 250 to 5600 cycles is divided into only three bands.

Room Noise in Residence Locations.—Measurements were made in 96 residence locations.

Figs. 6 and 7 show distribution curves for these measurements. Compared with the corresponding measurements made in business

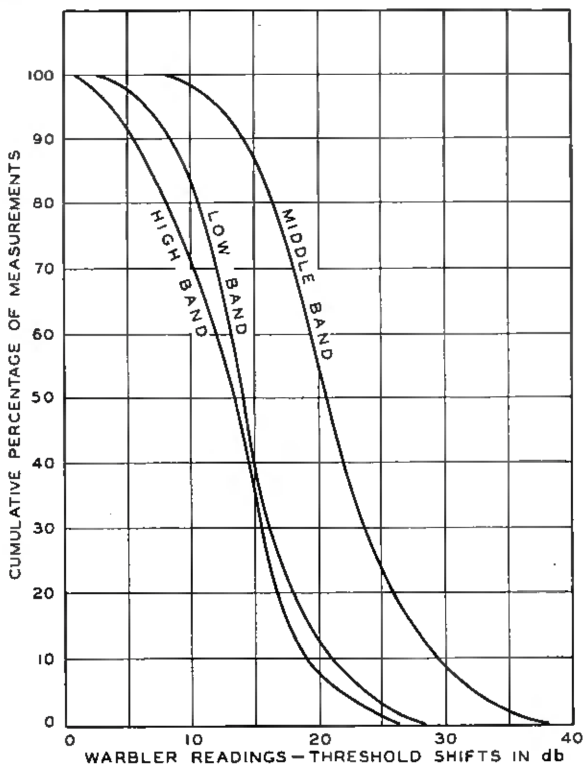


Fig. 7—Results of measurements of noise in residence locations by masking method.

locations it is apparent that the room noises encountered in residences were not only much smaller in magnitude but also varied less in magnitude than business room noises. The average of the residence room noises is about 18 db less than the average of the business room noises, while the spread in residence room noise (using the 95 per cent and 5 per cent points on the curves as limits) is 20 db, compared to 40 db for business noise; the standard deviation of the residence measurements is 6 db, compared to 12 db for the business neasurements. Unlike the curves for business noise, the curves for residence noise are very symmetrical, showing similar distributions above and below the average values.

As shown on Figs. 6 and 7, the median residence room noise would produce a reading of 7 db on the meter, and a masking of 12 db on the high-frequency warbler tone, 20 db on the middle-frequency tone, and 13 db on the low-frequency tone. The average was about the same as the median.

The average of the room noises measured in single-family houses was practically the same as the average of the noises measured in apartments.

Averages of the residence room noise measurements obtained in towns of various sizes are shown in the following table:

Size of town	Masking of middle-frequency tone	Meter reading	Number of measurements
Class A (over 400,000 pop.)..........	20 db	7 db	33
Class B (100,000 to 400,000 pop.).....	20 "	8 "	14
Class C (10,000 to 100,000 pop.)......	23 "	6 "	37
Class D (less than 10,000 pop.)......	17 "	7 "	12

It will be observed from this table that the residence noises measured in large cities were no greater than those measured in smaller towns. A study of the data showed that 27 of the 33 measurements made in Class A towns were made in residences which would be classed as apartment houses. It is possible that the noise usually associated with big cities is confined chiefly to non-residential locations, and that apartments on side streets are no noisier than residences in smaller towns. It should be recalled, however, that the number of measurements in each class of town was very small. In any case the data tend to show that the difference between residence noise in the large city and that in the smaller town probably is not extremely large. The measurements for Class A cities were made chiefly in Manhattan and Brooklyn with a small number in Newark.

It was found, in a manner similar to that discussed above for business noise, that the average residence room noise was quite complex in frequency makeup, and apparently did not differ materially from the average business noise in the relative amplitudes of low and high frequencies.

Comparison of Room Noise in Winter and Summer.—Forty locations were visited both in summer and in winter and the data compared. It was found that both business and residence noises were somewhat greater in summer than in winter, the average difference being 4 or 5 db. The spread in values obtained under summer conditions was less than that found for the winter conditions. This was because the noises which showed the least magnitude, when measured in winter, were found to be higher under summer conditions, while the highest noises measured failed to show an appreciable change with season. These highest noises were largely caused by indoor machinery, and would not be appreciably modified by outside sources.

The average frequency composition of the noises measured under both summer and winter conditions seemed to remain about the same as far as could be determined.

Selection of Typical Room Noise and Its Reproduction

The data obtained have been used in determining the characteristics of a typical room noise to be recorded on a phonograph record and reproduced for use in laboratory tests.

Since the data revealed no difference between the average frequency composition of great and small noises, it has been possible to choose a single recorded noise and to vary merely the amplitude of the reproduced noise, keeping its frequency makeup constant.

The recording and reproduction of such a noise have presented problems, particularly from the point of view of naturalness. It has been found difficult to reproduce a noise by simple means in such a way as to give the illusion that the noise is real, not artificial. The requirements for reproducing a noise which will be typical in its effect on the intelligibility of speech transmitted over telephone circuit are, however, considerably less severe than those for obtaining naturalness. Three main factors seem to be involved in the problem. In the first place, room noises often contain high-frequency components, undoubtedly including some extremely high frequencies. These components, while they are generally of low energy content, seem to contribute substantially to the naturalness of the sounds. The effect of these components on the intelligibility of speech transmitted over a telephone circuit would, however, be much less than their contribution to the naturalness of the noise, since the transmitted speech is generally limited to a band of not more than 3000 cycles. The frequency band transmitted by the recording and reproducing system was nearly twice this amount, being limited both by the mechanical characteristics of the apparatus and by the unavoidable noise generated in this apparatus, the amount of this noise increasing as the band width increases. Second, room noises emanate from a considerable number of sources located in different positions, so that in order to reproduce them with complete fidelity

LIST OF TOWNS WHERE ROOM NOISE SURVEY MEASUREMENTS WERE MADE

Size of town	Name of town	Number of measurements	
		Business	Residential
Class A (over 400,000 pop.)	Brooklyn, N. Y.	7	18
	Manhattan, N. Y.	32	13
	Newark, N. J.	0	2
	Total	39	33
Class B (100,000 to 400,000 pop.)	Jamaica, N. Y.	11	11
	Yonkers, N. Y.	7	3
	Total	18	14
Class C (10,000 to 100,000 pop.)	Bloomfield, N. J.	0	2
	East Orange, N. J.	7	9
	Flushing, N. Y.	0	5
	Harrison, N. J.	3	0
	Kearny, N. J.	6	0
	Maplewood, N. J.	6	10
	Milburn, N. J.	4	1
	Mt. Vernon, N. Y.	6	4
	New Rochelle, N. Y.	0	2
	Orange, N. J.	4	0
	Summit, N. J.	2	2
	West Orange N. J.	3	2
	Total	41	37
Class D (less than 100,000 pop.)	Hollis, N. Y.	6	7
	Madison, N. J.	0	3
	Pelham, N. Y.	3	1
	Richmond Hill, N. Y.	2	1
	Total	11	12
	Grand total	109	96

each source must be reproduced separately in its own position. On account of binaural effects in hearing, the proper locating of sources seems to have a considerable effect on naturalness. The most practical method of securing an approximation to this effect in the reproduced noise is to dispose a number of loudspeakers in different places in the room, chosen by test so that false directional effects are avoided. Third, the effects of reverberation must be considered. A noise picked up in a highly reverberant room, and reproduced in another highly reverberant room, would have in it two sets of reverberations. The best method of taking care of this seems to be to make artificial adjustments in the reverberation in the two rooms. Finally, there is a residual effect due to the fact that a person experiencing an actual noise is aided in his recognition of the noise by visual and other factors enabling him to refer it easily to its source; these are, of course, not present when the sound is reproduced.

Conclusions and Acknowledgment

While a certain amount of work on room noise conditions in telephone locations had been previously carried out, this survey represents a considerable advance in knowledge of room noise magnitudes. It provides data for work on the effects of noise on telephone transmission as well as furnishing certain information of wider interest. The methods of measurement employed, when further developed in the light of the experience gained in this work, should prove valuable in other room noise investigations.

The authors wish to acknowledge the work of Messrs. J. W. Whittington and R. E. Philipson of the National Electric Light Association and Messrs. J. M. Barstow and R. S. Tucker of the American Telephone and Telegraph Company, in designing and building the room noise meter and in carrying out the survey.

Contemporary Advances in Physics, XXI

Interception and Scattering of Electrons and Ions

By KARL K. DARROW

This article deals with a couple of aspects of one of the amplest questions of modern experimental physics: the question of what happens when an electron (more generally, an electron or a proton or a charged atom of any kind) collides with an atom or a molecule. It is well known, of course, that if the electron has energy enough, it may excite or ionize the atom. There are many different modes of excitation, and often several of ionization; the variety of possibilities is wide. If any of them occurs at an encounter, the electron loses energy and speed, and may suffer a change in the direction of its motion—a "scattering," as this is called. Even if it loses no measurable amount of energy at a collision, it may be "scattered," that is to say, deflected. The scattering and the energy-losses of the electrons are studied both on their own account, and because of the light they shed on what is happening to the atoms.[1]

IMAGINE a stream of electrons projected, all with known and uniform velocity and along the same direction, into a rarefied gas. Perhaps it ionizes the gas; if so, positive ions appear, and one may detect them and identify them and count them in any of various ways, without concerning oneself about the destiny of the ionizing electrons. Or perhaps it excites the gas without producing ions; if so, the atoms (or molecules) send forth light, and one may detect the excitation and identify the manner and measure the likelihood thereof, without paying any attention to the corpuscles responsible. Nevertheless, these corpuscles also must have been affected; they must have given up some at least of their kinetic energy, and if they still retain some motion, it is probably no longer in the same direction as at first.

If there is ionization or excitation of the gas, there should be electrons wandering off obliquely from the stream, and moving more slowly than when they entered the gas; in technical language, there should occur "scattering with loss of energy." Even if the incoming corpuscles are moving too slowly to ionize or excite, there might be—and there are—electrons wandering off obliquely with practically undiminished speed; they have suffered "elastic impacts" with atoms or molecules, they have been deflected merely, or "scattered without loss of energy." And even if the incoming corpuscles are moving fast enough to ionize or excite, some may be scattered with undiminished speed while others are spending some of their energy in these operations. Also, some of

[1] This article, in somewhat altered form, is to appear as a chapter in a forthcoming book on ionization and conduction in gases.

the electrons may adhere to atoms or molecules.[2] One sees that there are several items of knowledge about the effects of the gas upon the electron-stream, which cannot be discovered by studying the inverse effects of the electron-stream upon the gas, the formation of ions or of excited atoms; it is necessary to observe the stream itself. Even our knowledge of the latter effects may be improved by examining the former. These are the purposes of the experiments which I shall now describe.

The first (and the most) of these experiments may seem rather paradoxical, in view of what I have just been saying; for they are experiments not on electrons of the stream, but on the *absence* of electrons from the stream. One sends a beam of these corpuscles (or, it may be, a beam of protons or of once-ionized potassium atoms) across a stratum of gas, measures the number which go in and the number in the emerging beam, and puts down the value of the difference as the number which "vanish from the beam." "Vanish" is a good word in this connection; it is not meant, of course, that the missing corpuscles and their charges literally cease to be; it is meant simply that they do not belong among those which go straight through with undiminished speed and undeflected path, as though the gas were not there. I will say that they have been "intercepted," for this is a word which does not imply any choice among the varied possibilities of stoppage, adhesion, and deflection with or without loss of energy. Experiments on interception of fast electrons—up to 30,000 equivalent volts—were first performed at the beginning of this century; but the earliest accurate work on slow electrons—say 50 equivalent volts and downward—is only ten years old.

The results of these experiments are very striking; but of course they yield only a small part of what is wanted. We want to know what becomes of the "vanished" electrons, which way they have gone and with what residual speed—the total number and the distribution-in-direction of those which have been scattered without loss of energy; the total number and the distribution-in-energy and the distribution-in-direction of those which have ionized or excited the atoms which they struck; and the number of those which have stuck to atoms, if such there be.

Such information, as anyone would foresee, is harder to acquire. Of the distribution-in-direction of the scattered electrons, nothing was known four years ago; and what in this last quadrennium has been ex-

[2] Some gases being monatomic and others not, it is necessary to say "atoms or molecules" when making general statements, if one wishes to be exact; but in the following pages I shall often use either word by itself, even when the statement in which the term occurs is meant to apply to gases of both kinds.

plored is very little, compared with what remains. As for the distribution-in-energy, the first step in determining it was taken some sixteen years ago. It was indeed a great step; for it led to the discovery of the process of excitation, the transfer of energy from moving electrons to atoms which shift these latter from their normal condition into one or another of their "excited" states.[3] But it was only a beginning; the method had to be much modified and refined, to make it capable of finding the answers to such questions as I have phrased above; and the modifications were scarcely even imagined as lately as four years ago. Hence the reader must not expect to be introduced to a very great body of systematized knowledge. As for experiments on protons and on other kinds of charged atoms, they too are all extremely recent.

I begin with the experiments on interception of electrons.

Suppose then that a beam of electrons is sent across a tube, having first been limited by a sequence of slits or holes so that it has a definite contour, like the beam of a searchlight, which it retains all the way across the tube if there is vacuum. Further, suppose that on the far side of the tube there is a collector just large enough to swallow up the entire beam so long as it does not spread, but no larger; or alternatively, a collector covered by a screen pierced with a hole just large enough to let the beam, or a fraction of it, pass through. Even so, the result may depend on the diameter of the beam in a way which the reader will see for himself later on; it is best to think of a very narrow pencil of corpuscles.[4]

When a gas is introduced into the tube, the current into the collector will decrease. The decrease will be proportional to the density of the gas, so long as this is not too great; and it will be possible to define a "cross-section of the molecule for interception of electrons" in the same general way as is the custom in many other fields. Which is to say: denote by dx the distance which the beam traverses through the gas; by Q the number of electrons which enter the gas per unit time, hence by Qe the amount of charge which in unit time would arrive at the collector were the gas away; by Re the amount which in unit time does actually arrive at the collector; by N the number of molecules per unit volume: then the cross-section in question—call it σ—will be defined

[3] That is to say, the discovery by experiment; it had been predicted by Bohr (for the history of these matters, see for instance my "Introduction to Contemporary Physics," Chapter VIII).

[4] Of course this is an ideal which can never be perfectly realized. No matter how many diaphragms may be set up in a row to narrow and sharpen the beam, there will always be transverse motions of the electrons, relatively more important the smaller the forward velocity is made. Moreover the mutual repulsion of the corpuscles will tend to widen out the beam by driving its members apart. This is one of the reasons why experiments in this field were first performed on fast electrons, then extended to smaller and ever smaller velocities as time went on and technique was improved.

by the equation:

$$Q - R = \sigma NQdx. \tag{1}$$

The greater it is, the greater the fraction of the number of incident electrons which are intercepted, the greater the probability of being intercepted for any one electron; it is thus a measure of a probability of likelihood, the "likelihood of interception."

The quantity $(Qe - Re)$ is "missing current"; it is the amount by which the current to the collector drops off, when N molecules per unit volume are introduced into the tube. Nothing has yet been said, nothing has even been implied, about the fate of this lost current and about the missing electrons which presumably bore it into the gas. I have, in fact, been using the very neutral word "interception" so as to evade all implications in excess of what the data say, which is, that some of the electrons fail to persist in the beam. Not to suppose that they have been annihilated, there are at least two conceivable things which may have happened to them. They may have made elastic impacts against molecules, bouncing off in new directions, and being thus deflected out of the beam without suffering much change in speed. Or, they may have struck and stuck to molecules moving in other directions than that of the beam. Other possibilities are thinkable; but these are enough to hold in mind for the present.[5]

(The word "absorption" is used by some, especially by Germans, in the sense for which I here use "interception." It seems to me to convey unwanted implications, but there may be differences of opinion on this point. Much the same problem of language occurs in optics. Usually the term "absorption of light" means in practice "departure of photons from a beam of light" irrespective of whether they are actually swallowed up by atoms, or deflected without any loss of energy; but it is rather common nowadays, especially in treating of X-rays, to use "absorption" for the former mode of disappearance only, and "scattering" for the latter. Since it is necessary now to distinguish two kinds of scattering of photons, the complications are beginning to rival those of electronics.)

Adopting either the elastic-impact idea or the adhesion idea, we may visualize this quantity σ in a familiar way. We may conceive of the molecules, for this purpose and for this purpose only, as spheres so constituted that when an electron touches one of them it sticks—or else rebounds, whichever theory we are using. The value of σ is then

[5] Lenard reviewed a number of possibilities, and considered ways of distinguishing them in his brochure *Quantitatives über Kathodenstrahlen.* He made a peculiar distinction between reflection of electrons from molecules, and small deviations of electrons by molecules; it seems to have been suggested by his work on very fast corpuscles.

the value which must be assigned to the cross-section of these spheres, in order to make the calculated values of "missing current" agree with the observed ones.

An elastic-sphere model is also used in the kinetic theory of gases: one visualizes a gas as a flock of spheres, and for their cross-section one chooses the particular numerical value which, when inserted into the kinetic-theory formula for the viscosity of that gas, gives a figure agreeing with the measured viscosity. This is the so-called "gas-kinetic cross-section," which I will denote by σ_0. One should not expect it to be identical with the quantity σ which has just been defined, nor be surprised at finding differences—even differences in order-of-magnitude—between the two. The elastic-sphere model is good for many purposes; but it has its limitations.

The ratio of σ to σ_0 is occasionally used instead of σ as a measure of the likelihood of interception. Much more frequent of usage [6] is the product of this σ by N_1, the number of molecules in a cc. of a gas at 0° Centigrade and one mm. Hg ($N_1 = 3.56 \cdot 10^{16}$). So also is the reciprocal of $N_1\sigma$, the so-called "mean free path" of the electrons under the stated conditions, zero Centigrade and one millimetre pressure. It should be called "mean free path for interception," but the qualifying word is usually left out. The reciprocals of $N\sigma_0$ and of $N_1\sigma_0$ are also frequently used as standards of comparison, sometimes with the name "gas-kinetic mean free path." Since the concept of mean free path is used quite often in stating the results of these experiments, I will give the reasoning whence follow at once its definition, and its relation to the quantity σ.

Returning to equation (1), rewrite it thus:

$$dQ = -\sigma NQdx. \tag{2}$$

Here the quantity $(Q - R)$, the number of electrons lost per unit time from the beam between the planes x and $x + dx$, is written as $-dQ$, the negative of the difference between the numbers which in unit time cross the planes $x + dx$ and x respectively. Integrating, we get:

$$Q_1/Q_2 = \exp[-N\sigma(x_1 - x_2)] \tag{3}$$

for the ratio between the numbers crossing any two planes separated by the distance $(x_1 - x_2)$.

For simplicity, suppose that it is at the plane $x = 0$ that the corpuscles enter the gas, and denote by Q_0 the number entering per unit time. Then the number $Q(x)$ reaching any plane x is this:

$$Q(x) = Q_0\exp(-N\sigma x), \tag{4}$$

[6] Especially by Ramsauer, by Brüche, and other Germans generally.

and this is the number of electron-paths which extend unintercepted through the distance x measured from the plane $x = 0$. Out of these, the number which are terminated between the planes x and $x + dx$ is this:

$$- dQ = N\sigma Q dx = N\sigma Q_0 \exp(- N\sigma x) dx \tag{5}$$

and this is the number of electron-paths which, measured from the plane $x = 0$ to their termini, have lengths between x and $x + dx$. Multiply it by x, and you have the sum of the lengths of all these paths. Integrate this product $xN\sigma Q dx$ from $x = 0$ to $x = \infty$, and one has the total length of the unintercepted paths of all the Q_0 electrons; divide the integral by Q_0, and one has their mean length l, the "mean free path for interception":

$$l = \int N\sigma x . \exp(- N\sigma x) dx = (N\sigma)^{-1}, \tag{6}$$

and this is the reason for giving the name "mean free path" to the reciprocal of $N\sigma$. As for the reciprocal of $N_1\sigma$, it is the value of mean free path at zero Centigrade and one millimetre pressure; it may be denoted by l_1:

$$l_1 = (N_1\sigma)^{-1}. \tag{7}$$

Though everywhere along this train of reasoning the paths were supposed to be measured from the plane $x = 0$ where I said that the corpuscles entered the gas, the result is not restricted. No matter where the plane, from which the paths are measured (so long as it lies in the gas), the mean of their lengths from that plane to their various terminations has the same value $(N\sigma)^{-1}$. It follows that if these so-called "interceptions" are elastic collisions, from which the electrons rebound with practically the same value of speed as they had beforehand, the mean-free-path from one collision to the next should likewise be equal to $(N\sigma)^{-1}$. But this is a deduction which we can hardly hope to check by simple experiments on a beam of electrons, since after its first collision a corpuscle quits the beam. Of course, if the "interceptions" are adhesions of electrons to molecules, there is no sense in making this extension.

Something must now be said about the relation between these "mean free paths" of electrons, and the quantity called the "mean free path of the molecules of the gas." Even if σ were equal to the gas-kinetic cross-section σ_0—even if the molecules behaved towards one another as elastic spheres of radius $\sigma_0/\sqrt{\pi}$, and towards electrons as elastic spheres of the *same* radius $\sigma_0/\sqrt{\pi}$—the mean free path of a molecule between collisions with its mates would not be the same as the mean free path of an electron between collisions with molecules. For, conceiving an electron as a mere point (as we are regularly doing); if it is to hit a molecule M, the direction of the motion of its centre, which is itself, must be pointed towards M. But if instead of an electron it is another molecule M' which is to hit M, the sufficient condition is that the direction of motion of the centre of M' should be pointed towards a sphere of twice the radius, four times the cross-section of M or M'. Merely on this account, the mean free path of a molecule among similar molecules should be only one-fourth as great as it would be, if the moving particle were shrunk to a point. Another allowance must be made for the fact that all the molecules of a gas are in motion, although their motions are so slow that relatively to a free electron they may be viewed as stationary. I shall not derive the formula for this latter allowance; but the net result of both may be expressed in this way: the ratio of the mean-free-paths of an electron and a molecule, in a given gas, should be $4\sqrt{2}$ if the ratio σ_0/σ of the gas-kinetic cross-section to the cross-section for electron-impact were equal to unity. This result however is based on so very specific and in part fallacious assumptions, that I should not have treated it at such length, but for the fact that it is more or less the custom to divide the mean-free-path for interception of electrons by $4\sqrt{2}$, and compare the result with the mean free path of molecules as the kinetic theory of gases supplies it from the viscosity of the gas. This amounts in practice to using still another measure of the "likelihood of interception": to wit, the ratio $4\sqrt{2}\sigma_0/\sigma$.

Doubtless it seems that I have spent excessive time in talking of these measures of the likelihood of interception. There is however, an unimpeachable reason for dwelling on the topic: the likelihood is the only thing which *can* be measured. There are no critical "intercepting potentials" like ionizing or resonance potentials; there are no specific energy-values of the impinging electrons at which interception abruptly starts; there is nothing in the nature of sudden "onset" to be detected. The probability of the effect is the physical reality, it *is* the effect; and its relation to the speed of the electrons is the only quality available for study. Unless we know precisely how we are defining it, we know nothing. Also we shall soon be looking at experiments which yield values for quantities much like the σ of interception, and yet not quite the same; it will be necessary to discriminate with great care. But let us first consider the direct experiments of the type which I have been presuming.

The experiments would certainly be simplest, if the electrons were dashing along with enormous forward speeds, corresponding to vis viva of the order of thousands of equivalent volts; for then their transverse speeds would probably be negligible, the beam would have little tendency to spread other than that for which the gas itself is accountable. As early as 1894, Lenard did such experiments with 30-kilovolt electrons (cathode-rays which had emerged from a discharge-tube through a window of thin metal foil). A few years later, he added data obtained with slower electrons, and the work was continued by Becker and by Silbermann.[7] For the 30-kilovolt corpuscles, the cross-section σ is very much smaller than the gas-kinetic σ_0—only a few per cent as great. As the energy of the electrons is decreased, σ rises towards σ_0. All this is illustrated in Fig. 1.

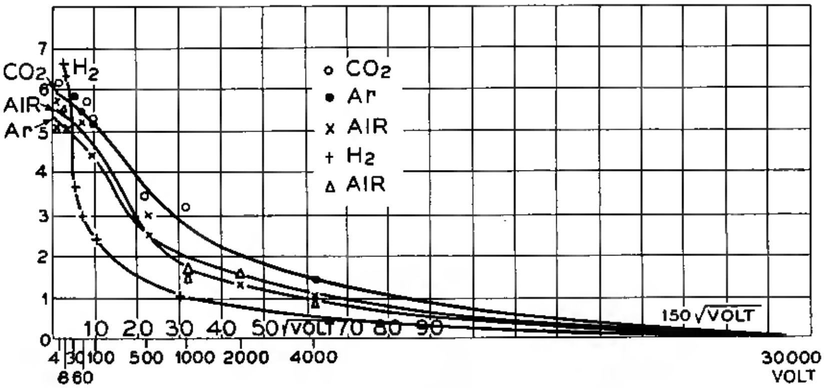


Fig. 1—Cross-section for interception of electrons, plotted for various gases over a wide range of electron-speeds. (P. Lenard, *Annalen der Physik.*)

Obviously, these are not experiments in which the corpuscles are unable to ionize or to excite; quite the contrary. One might be tempted to rush to the other extreme, and guess that all of the electrons missing from the beam have effected ionizations, that the quantity σ of Lenard's experiments is a measure of *likelihood of ionization;* but in the present state of knowledge, this would be going too far. It is important, however, that for the *fast* electrons σ turns out to be proportional to the atomic number of the gas, if this is monatomic; and to the sum of the atomic numbers of the atoms constituting the molecule, if the gas is diatomic or triatomic. This sort of rule suggests that interception of fast-flying corpuscles is due either to the nuclei of the atoms, or to some action of the bound electrons of the atoms in which they all are equally potent, however tightly or loosely they are bound.[8]

[7] References are given at the end of the paper.

[8] Silbermann (dissertation, Heidelberg, 1910) and A. Becker investigated a large number of organic compounds, using fast electrons; and they found that the σ of each of these molecules (some of them composed of five to twenty atoms) could be predicted accurately, simply by adding together the values of σ for the constituent atoms as determined by experiments on simpler gases.

In his latest experiments Lenard went down to electron-speeds so low that he was working near, or even below, onset of ionization; but instead of dwelling on these, I will take up the experiments which are designed primarily for the study of slow electrons—experiments, the first of which were made in Lenard's own laboratory by his associates, H. F. Mayer and C. Ramsauer. They tried in somewhat different ways to realize in practice the ideal scheme of apparatus which in the last few pages I have been taking for granted. Before depicting their devices, I revert for a moment to equation (4) above.

Suppose that the collector is located at the distance x from the aperture where the beam of Q_0 electrons per second enters the gas: the number Q of electrons and the charge Qe reaching it in a second should conform to the epuations:

$$Q = Q_0 \exp(-N\sigma x), \tag{8}$$

$$\log(Qe) = \log(Q_0 e) - N\sigma x. \tag{9}$$

Thus on plotting the logarithm of the collector-current against x one should get a straight-line graph, and *the slope of the line should give the value of* $N\sigma$, therefore the value of σ when the density of the gas is known. It should suffice to slide the collector along the direction of the beam, and plot the logarithm of the current against the distance through which it has slid, measured from any arbitrary zero. Or alternatively, one might keep the collector stationary at some fixed distance x_0 from the point of entry, and vary the density of the gas. Plotting logarithm of current against N, one should get a straight-line graph, and the slope of the line should give the value of $x_0\sigma$. It is then unnecessary to worry about the possible presence of residual vapors of unvarying density not recorded by the pressure-gauge, for they would contribute only an additive term to N, not affecting the slope of the line.

It must be granted that the interpretation of the data is seldom quite so simple. The quantity Q_0 may vary with the density of the gas; some observers determine the total emission of electrons from the source (or something which they take to be proportional thereto) for each value of N separately, and plot the logarithm of the ratio of Qe to this latter as function of N. The beam may diverge even in a vacuum, or what the experimenter takes to be a vacuum; some measure Qe twice for each value of x, once with a vacuum and once with gas, and plot the logarithm of the ratio of the two as function of x. Mayer varied both N and x and combined the results in the hope of thus eliminating the effects both of the divergence of the beam in a vacuum, and of residual gases.[9]

[9] It is not always easy to make out from a paper just what procedure the observer has followed, nor how far he has traced the curves of Q vs. x or Q vs. N.

(I pause to point out the relation between formula (8) and the earlier formula (1). The first two terms in the Taylor expansion of e^{-y} being $(1 - y)$, one may write instead of (8),

$$Q = Q_0 - Q_0 N \sigma x, \tag{10}$$

which is formula (1).)

The experiment of Ramsauer acquired instant fame, because the data which he got with argon were of a nature totally unexpected,[10] and caused immense surprise.

His apparatus is a metal box, partitioned into chambers, the party-walls of which are pierced with slits delimiting a narrow path curved in the form of a circle. Details are different in the different boxes which he used at various times, and in those which several other physicists—

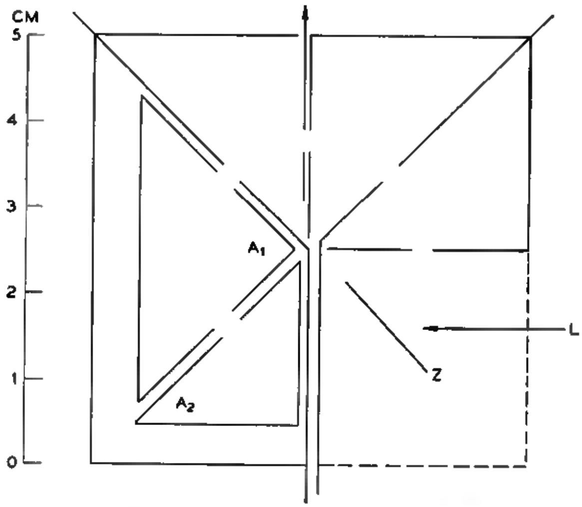


Fig. 2—Ramsauer's apparatus for measuring the cross-section for interception. (*Physikal. Zeitschrift.*)

Maxwell, Brode, and T. J. Jones among them—made after the pattern of his; but a fair idea of all is given by Fig. 2. Photoelectrons [11] escape from the metal plate at Z, and are accelerated to the speed desired by a potential-rise from Z to the walls of the box. The electron-beam consists of corpuscles sweeping around in circles so centred and so proportioned that they pass through all the slits. It is of course a magnetic

[10] This was first disclosed in Mayer's paper from the same laboratory, but Mayer yields to Ramsauer the credit of having noticed it sooner.

[11] Some of the American physicists used thermionic electrons instead; in certain experiments the filament replacing Z is encased in a coaxial cylinder with a narrow slit, from which electrons escape after receiving the energy corresponding to the potential-rise from filament to cylinder.

field, perpendicular to the plane of the drawing, which causes them to revolve so. This field limits the speeds of the electrons in the beam to a certain range, which may be narrower than the range over which the speeds of the electrons emerging from Z are spread.[12] It has a second purpose, which we shall see directly. The electrons are collected in either of the chambers A_1 and A_2, alternate use of which enables the observer to vary the distance x figuring in equation (8). Mostly, however, it is the density of the gas which is varied.

If there is gas in the box, electrons which by striking molecules are deflected even a little go to one of the partitions, and vanish from the beam. So also do electrons which strike molecules and stick to them; even were one to adhere to a particle of molecular size which happened to be moving in the direction tangent to the beam, the resulting massive ion would travel in a path with a different curvature, and fall against a wall. Again, if an electron should suffer a loss of kinetic energy without losing its freedom or its direction of motion, it too would be eliminated from the beam; its path would be more curved after the impact than before, and it would miss the slits. Again, if by an ionizing impact a fresh electron should be ejected from a molecule, it would be moving more slowly than those of the primary beam, and could not stay with them. Finally, if an electron were bounced out of the beam by an elastic impact with a molecule, it could not be bounced back in in again by fewer than two additional collisions, and these it would be most unlikely to make. Clearly, the quantity σ of which this apparatus gives the value is very stringently defined!

Suppose now the magnetic field omitted, and the slits and chambers of the device of Fig. 2 all arranged along one straight line. This gives another well-known type of apparatus. The first of this kind was the one which Mayer used, but I choose for representing here the one designed by T. J. Jones (Fig. 3). The electron-beam, consisting of electrons which emerge from the filament F and are accelerated by a rise of potential between F and the screen D_2, is shaped by the sequence of slits which is shown in the sketch, and enters the chamber B through

[12] Ordinarily the magnetic field is so adjusted, that the range of electron-speeds aforesaid is centred at the value which is the most probable speed of the electrons. The corpuscles which have come out of the source and have been accelerated to the first slit have a certain distribution-in-speed with a maximum at some special value, say u_0; the range selected by the magnetic field has u_0 as its center. However Ramsauer and Kollath found that they could not do this with electrons having less energy than 0.45 equivalent volt; they then adjusted the field so as to select ranges of speed lower than that which included the maximum, and were able to extend their experiments as far down as to electrons of 0.16 equivalent volt. The use of the magnetic field has the incidental advantage, that the values of electron-speed deduced from its value and from the radii of the circular orbits are correct even if there are contact-potential differences in the box—something which cannot be said for those deduced from the voltage applied between box and filament.

the slit S_1. The part of the beam which is not intercepted goes on into the chamber C. In the notation of equation (8), the current Qe is that reported by the galvanometer G_2, the current Q_0e is the sum of those reported by G_2 and the other galvanometer G_1; the distance x is measured from the slit S_1 to the pair of slits S_2S_3, and therefore has a constant value x_0, while N is varied by admitting or withdrawing gas. In certain experiments of Brode and in those of Rusch, the electron-beam passes through a metal-walled channel—as though the slits S_1 and S_2 were circular apertures, the opposite ends of a metal tube. In the similar devices of Mayer and of Maxwell, the distance x is varied by sliding the chamber C back and forth, thus lengthening or shortening the chamber B; for which purpose, B and C are made of lengths of tubing of which the latter telescopes into the former.

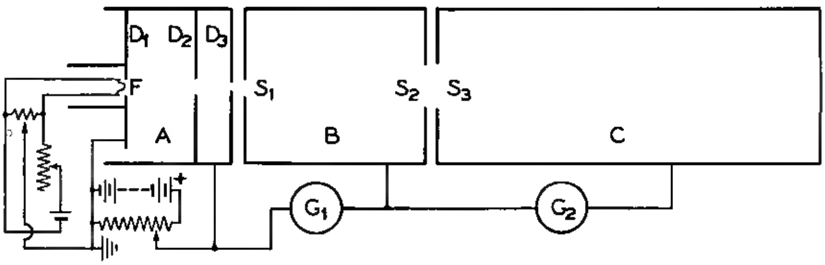


Fig. 3—Apparatus for measuring cross-section for interception, of the "Mayer" type. (T. J. Jones, *Physical Review.*)

Comparing this now with Ramsauer's device, one sees that in either apparatus, electrons which are considerably deflected in collisions fail to reach the collector. However, among the various effects which were listed in the last paragraph but one, which a molecule might conceivably produce, there are some which would cause electrons to quit the beam as defined in the Ramsauer apparatus, but not as defined in Mayer's,—for instance, a corpuscle which had lost much of its speed without suffering change of direction of motion would go right on into C of Fig. 3, as though it had had no encounter at all, while in the Ramsauer scheme it would figure among the missing. The quantity σ determined by a scheme of Mayer's kind is thus essentially different from the σ determined by one of Ramsauer's kind; and indeed any change in detail of either, any widening of the slits for example, should involve a change in the meaning of the σ which the experiment reveals—to each apparatus, its own cross-section of the molecule! Thus by applying a retarding-potential between the slits S_2 and S_3 of Fig. 3, one might withhold from the collector such electrons as had suffered much loss in speed but little deviation,[13] and the σ measured by the so-modified device would approach more nearly to that of Ramsauer's. Modifications like this might conceivably entail enormous changes of the relevant cross-section. But it happens that they do not—a very important fact[14] in itself, illustrated by Fig. 7 (for mercury vapor); and therefore the results of all the work can be summarized *en bloc*.

I begin with the surprising fact, of which the disclosure excited so wide an interest in this field. Argon was the gas on which the dis-

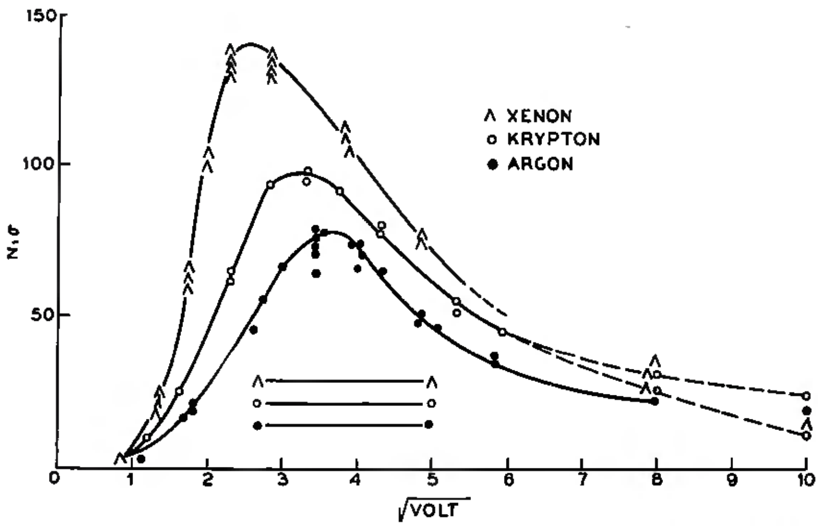


Fig. 4—Cross-sections of atoms of three of the noble gases, determined by Ramsauer with the scheme of Figure 2. (*Physikal. Zeitschrift.*)

covery was made by Ramsauer, and checked by Mayer (and a little later by Townsend and Bailey, in an entirely different way; but since krypton and xenon show even more strikingly the "anomaly"—as it used to be called, and should be called no longer—I present the curves for all three of these gases together (Fig. 4). They are curves of $N_1\sigma$ (the reciprocal of l_1) versus the speed of the electrons.[15]

[13] This in effect was done by Mayer, whose collector was shaped like a cup covered over with a pair of parallel gauzes; he made the cup and the inner grid a volt or two more negative than the outer gauze. Notice that this is a scheme for detecting critical potentials for onset of inelastic impacts.

[14] Moreover, a distinctly puzzling fact, especially as sometimes the σ obtained by Ramsauer's method appears to be less than that obtained by Mayer's (see Fig. 7, and the paper of M. C. Green).

[15] In this department of electronics it is the custom to use as independent variable, not the kinetic energy of the electrons expressed in equivalent volts, but the speed of these corpuscles expressed either in centimetres-per-second, or (more commonly) in "square-roots-of-equivalent-volts," i.e. in a unit equal to $5.94 \cdot 10^7$ cm./sec. I do not know of any valid reason for this anomaly of usage, except insofar as it may be found that the corresponding curves for protons and other ions display similar features at equal speeds but not at equal energy-values.

Going towards lower speeds from the highest here represented, one sees that the curves are ascending. This is to be expected; they are the prolongations of the curves for fast electrons exemplified in Fig. 1; the cross-section for interception is steadily increasing as the corpuscle-speed diminishes, the atoms (these are monatomic gases) are bigger obstacles for slow electrons than for fast. But instead of always ascending, *they pass through maxima and sink as the electron-speed is further lowered.* This is the surprise. The atoms of argon, krypton and xenon are smaller obstacles for very slow electrons than for moder-

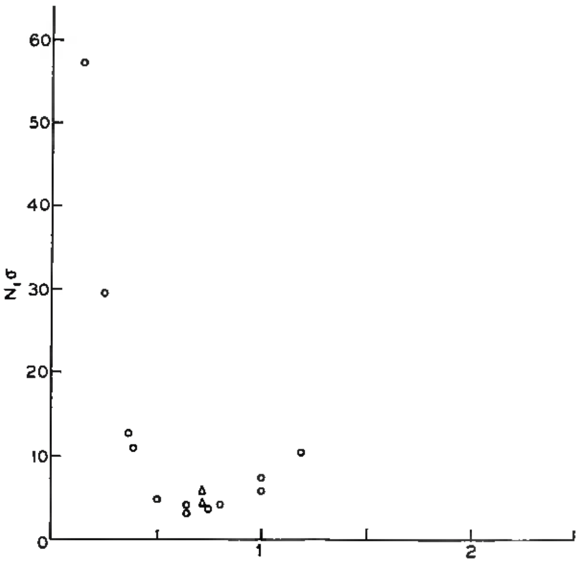


Fig. 5—Cross-section of xenon atom, for very low values of electron-energy. (C) Ramsauer, R. Kollath, *Annalen d. Physik.*)

ately slow ones. Below the maxima, there are minima and renewed ascents; these are very late discoveries of Ramsauer and Kollath, just confirmed by Normand and Brode; that of xenon is illustrated by Fig. 5.

The ordinates of the horizontal dashes in Fig. 4 are the values (multiplied by N_1) of σ_0, the gas-kinetic cross-sections of these three kinds of atoms. One sees that σ is several times greater than σ_0 at the maximum of each curve, several times smaller than σ_0 at the minimum. It was thought formerly that σ_0 should be the limit which σ approaches as the electron-speed approaches zero, and Lenard's early results with fast electrons seemed to sustain the notion; it perished at the advent of these data.

For helium and neon, on the other hand, the cross-sections vary but little over the energy-range from 1 to 40 equivalent volts, though for each there is a gentle flattish maximum. Below one equivalent volt, σ for neon falls gradually and smoothly as far as the limit (0.16 e.v.) of Ramsauer and Kollath's data, while the curve for helium has a couple of wiggles. Normand and Brode find a minimum in the neon curve.

It seems odd to speak of methane (CH_4) among the monatomic gases; but its σ-curve is of the same kind as those of the three massive noble gases, displaying a sharp maximum [17] near 8 equivalent volts. One notices however that if the four electrons of the four H atoms were to join the four outer electrons of the C atom, they could form a shell of eight to simulate the closed outer shell of an inert-gas atoms. But the noble gas most closely simulated should be neon, and neon does not show the high sharp maximum.

The other monatomic gases are the vapors of the metals. As in the measurements of their resonance and ionizing potentials, they are difficult to handle. There is an extra difficulty, over and above those which beset the seeker after critical potentials: the measurer of σ must know the density of the gas, therefore the vapor-pressure of the metal with which he is working. But vapor-pressures change so rapidly with the temperature, that this must be determined very carefully. Even after one has determined the temperature with exceeding care, one may find on searching the literature that the density-vs-temperature curve has never been reliably determined. The data for mercury are by far the most abundant; there are a few for cadmium and zinc, which belong to the same column of the periodic table (Fig. 6). All three have similar cross-section curves, which except for a little hump near 40 equivalent volts rise steadily with fall of corpuscle-speed. I reproduce an additional curve for mercury (Fig. 7), for it illustrates the smallness of the difference between the σ of the Ramsauer method and that of the method of Mayer. The "method of Part I" is substantially Ramsauer's; the "method of Part II" involved the use of the tube shown in Fig. 3.

The four familiar members of the alkali-metal column were subjected to experiment by Brode; each of the four curves of $N_1\sigma$ versus electron-energy has a sharp maximum near 1 or 2 equivalent volts, and this maximum is astonishingly high—greater than 1000 for all four, almost

[17] R. B. Brode (*P. R.* (2), **25**, 636–644; 1925). Akesson is said to have discovered this maximum, even before Ramsauer's research (*Lund Arsskrift*, 1916). This paper of Akesson's is the outstanding example of an article which hundreds cite for one who has seen it. For some reason—the war, very likely—it was never published in any journal enjoying a wide circulation.

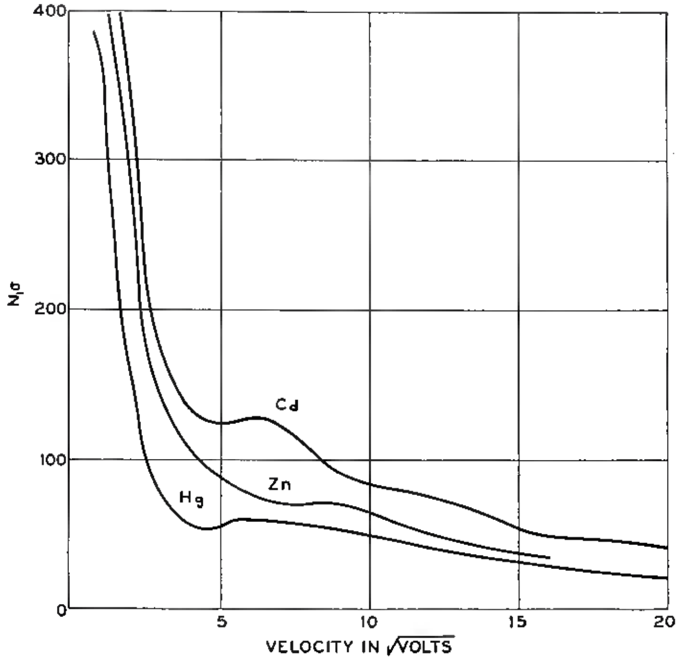


Fig. 6—Cross-sections of metals of the second column of the periodic table. (R. B. Brode; *Physical Review.*)

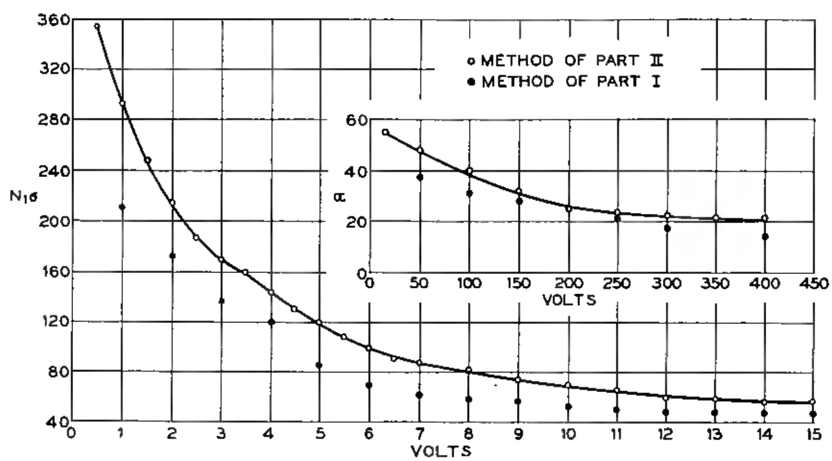


Fig. 7—Cross-section of mercury atom, determined by two methods, of the "Ramsauer" and "Mayer" type ("Part I" and "Part II") respectively. (T. J. Jones, *Physical Review.*)

2000 for cæsium. Dividing these values by N_1, we see that this means that for electrons of the corresponding speed, the atom is as great an obstacle as would be a sphere of cross-section 3.10^{-14} cm^2., of radius 10 Angstroms or more. Sizes such as these, when compared with the σ for ionization or with the gas-kinetic cross-section of the atom, are surprisingly large.

The molecules of the common molecular gases also yield curves with maxima, located however at energy-values lower than those at which the peaks for the noble gases stand. Brode in 1925 discovered a minor maximum for nitrogen at some 20 equivalent volts; but the really important ones lie much lower, those for hydrogen and nitrogen somewhere between 2 and 3. In Fig. 13 I show a curve for a trimolecular gas, carbon dioxide; Ramsauer has lately found that to the left of the point where this curve is cut off in the graph, it rises rapidly again. Oxygen shows a sharp minimum near 0.25. One sees that the so-called "anomaly" of argon is nothing anomalous at all; it is merely an example unusually conspicuous of a feature which atoms and molecules generally display.

Observations on the Scattered Electrons Themselves

Recall the classical experiment of Franck and Hertz—the one which led to the discovery of "inelastic impacts" of electrons against atoms, the discovery of the transfers of energy from electrons to atoms which result in excitation. There are three electrodes in a tube: a filament, the source of electrons—a grid, at a potential V volts higher than the filament—a plate beyond the grid, at a potential lower than this latter by a small and constant amount. The number of electrons arriving at the plate is plotted as function of V, and certain "breaks" are seen in the curve; these fix the location of the critical energy-values of the electrons, at which various modes of excitation first become possible.

Now this is not much different from the experimental method of Lenard, of Mayer and of others, in the work which I have just been citing; only, there is an important difference in aim; much more attention was paid by Franck and Hertz to the breaks, and the rôle of the retarding-potential between the plate and the grid. Franck and Hertz were studying electron-scattering with especial regard to the distribution-in-energy, and to the energy-losses, of the scattered corpuscles. In other work of theirs, they had auxiliary collectors off to one side from the grid, thus in effect studying scattering at large angles. It is further and more elaborate work of this sort which we have now to consider.

One scheme of apparatus was realized first by E. G. Dymond. I show in Fig. 8 a slight modification, set up by G. P. Harnwell. On the left, the electrons come from a filament contained in the cylinder *C*, are accelerated through a slit (1) in the cylinder and a second slit (2) in the front wall of a chamber beyond. (For arrangements of this sort, the word "electron-gun" has become the standard metaphor.) They meet the molecules in the center of the tube, and those which are scattered in directions passing through slits 3 and 4 go through these

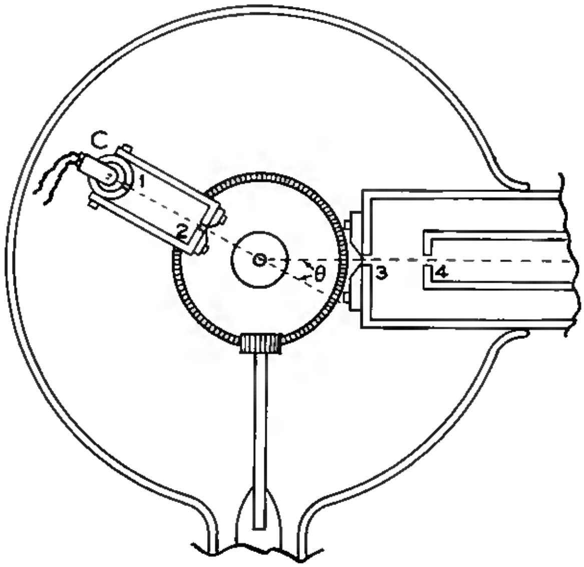


Fig. 8—Harnwell's apparatus for studying scattered electrons. (*Physical Review.*)

slits (unless they make further collisions) and onward into the "analyzer." The electron-gun can be revolved around the axis cutting the plane of the paper normally at 0, thus making it possible to study scattering as function of angle.

In the analyzer there is a region pervaded by a field, and beyond it a stationary collector, to which electrons go if they are suitably deflected by the field,—as in many a well-known type of analyzer used for identifying ions, for instance Aston's or Dempster's for finding isotopes. There, however, it is the charge-to-mass ratio of the ions which is computed from the strength of the deflecting field which sends them to the collector, their kinetic energy being taken as known. Here it is the energy of the electrons which is computed from the strength of the field, their charge-to-mass ratio being accepted as known. One plots the collector-current against the field-strength, and then, translating the later variable into electron-energy by means of the known relation, one gets the distribution-in-energy of the scattered electrons. (Incidentally, in Harnwell's apparatus and in that of MacMillen the deflecting field was electric; in Dymond's magnetic.)

Typical data of Harnwell's are shown in Fig. 9; these are distribution-in-energy curves for electrons scattered by helium, their initial energy having been 75 or 150 equivalent volts (curves on left and right, respectively). First, one sees that the great majority of those which go

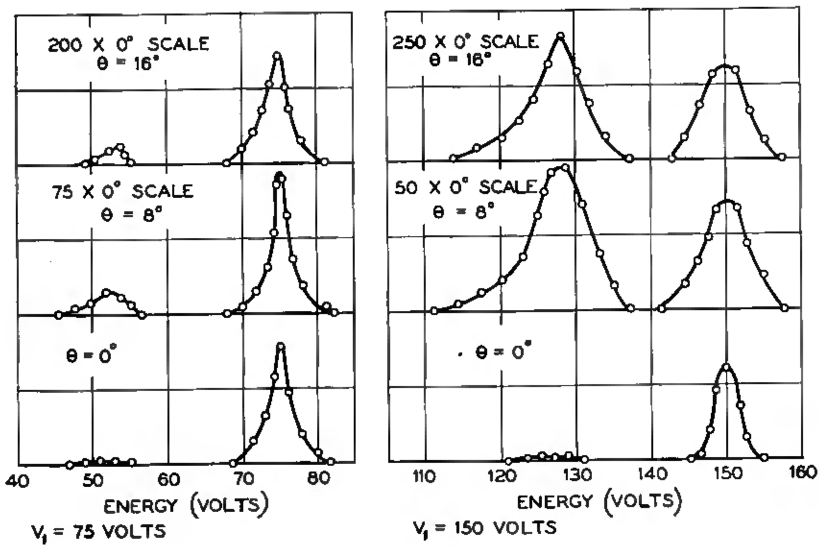


Fig. 9—Distribution-in-energy of electrons scattered from helium atoms, as determined by Harnwell. (*Physical Review.*)

through nearly undeflected retain their energy. Then, the electrons which are deflected through angles in the neighborhood of 8° are very much fewer (notice the change in the scale of the axis of ordinates) but among them, those which suffer a certain energy-loss are relatively more numerous. The electrons deflected at 16° are fewer yet, but among them those which suffer the same energy-loss are relatively plentiful. With 150-volt corpuscles one sees the same behavior, with differences in detail which I leave to be read from the curves. As for this peculiar value of energy-loss, its "mean value from a large number of observations is 22 volts"; one recalls that the resonance-potential of helium is 19.6, and that there are other critical potentials of inelastic impact ranging from this value upwards.

Curves obtained by Dymond and Watson, also for helium, appear in Fig. 10. The angle of deflection is about 10°, and the experiments were performed with 102-volt, with 226-volt and with 386-volt electrons. The double-pointed peaks bear witness to two distinct amounts of energy-loss frequently occurring; one maximum lies between 21 and 22 equivalent volts, the other is somewhat greater. Now the minimum amounts of energy which the helium atom can absorb, or in other words the energy-values of its lowest excited states, are 19.77 and 20.55 equivalent volts; and there are many others distributed between the upper of these and the ionizing-energy, which last is 24.5. It seems natural to

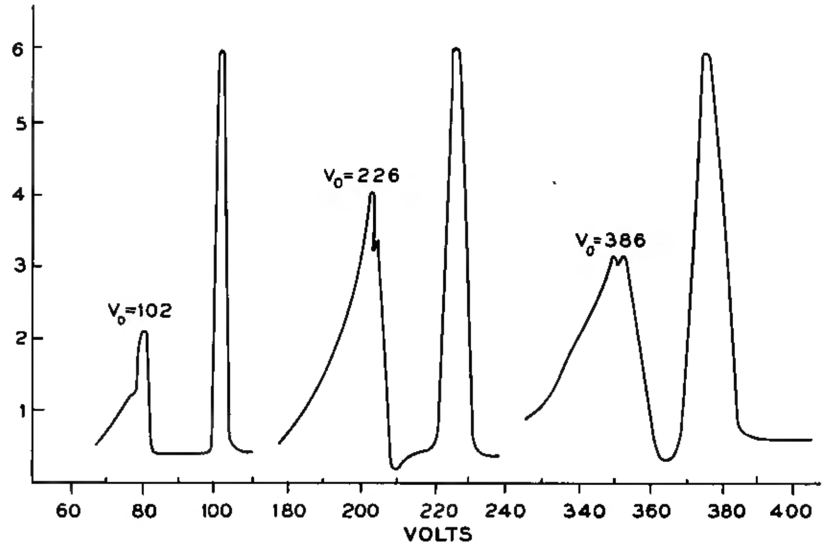


Fig. 10—Distribution-in-energy of electrons scattered at 10° from helium atoms, as determined by Dymond and Watson. (*Proc. Roy. Soc.*)

ascribe the less-displaced of the two points of the shifted peak to processes of excitation, the other to ionization; but one is tantalized to note that the peaks are not located quite accurately enough to settle this. Dymond and Watson find that as the angle of deflection is increased, the electrons scattered with undiminished energy take more and more the lead over those which contribute to the shifted peak.

Still other curves for helium can be seen in the article of MacMillen. Analyzing the 50-volt electrons scattered at 10°, he was able to plot a curve displaying no fewer than three distinct maxima, not counting the great one corresponding to the elastically-reflected corpuscles—evidence, therefore, of three distinct and distinctive energy-transfers. The most frequent of these he estimated as amounting to 21.50 equivalent volts, with a probable uncertainty of 0.15; the others are greater by 2.13 and by 2.95 equivalent volts. The accuracy is such that one may attempt to compare them with the energy-values of the excited states of the helium atom, but here there is a disappointment—the agreement is not so good as one would like. MacMillen is disposed to think that the value 21.50 should be corrected to 21.12 and the two others shifted equally, whereupon the first would agree accurately and the two others passably with the energy-values of known excited states. But these three do not comprise the two lowest of the excited states, the 19.77 and 20.55-volt levels which I mentioned above; it would then be necessary to assume that these are much less likely to be attained than three of those above them, which seems surprising but not impossible. At any rate it is evident that the time has arrived for experimenting in ways which permit of locating the shifted peaks with the highest accuracy possible.

Continuing with Harnwell's data: neon yielded a set of curves very like those of helium, except that the mean value of the energy-loss, deduced from the separation of the pair of peaks corresponding to the pairs in Fig. 9 amounts to some 18 equivalent volts; this is predictable. Molecular hydrogen displayed an energy-loss of 12.3 equivalent volts; but the most striking feature of the curves is the prominence of the peak formed by electrons which have lost that amount of energy—indeed, with 75-volt electrons, even those which go through nearly undeflected include a larger proportion of such, than of corpuscles which have retained their capital intact. Harnwell made measurements on nitrogen also, and on a mixture of molecular with atomic hydrogen, this being supplied from a discharge-tube in operation; and I reproduce as Fig. 11 his graph showing the distribution-in-angle of electrons scattered *without* loss of energy, the hollow circles relating to molecular hydrogen, the dots to the mixture.[18] The corresponding curves for electrons scattered *with* loss of energy lie very close together, and the one marked C in Fig. 11 represents them both. MacMillen traced similar but not exactly concordant curves for hydrogen, and others for helium and for argon. In his data also, electrons scattered through small angles predominate more and more, the greater their initial speed; and it is rather surprising to find how large a proportion of the corpuscles which have lost energy to helium atoms continue nevertheless with little or no deflection.

Somewhat earlier, Jones and Whiddington had studied the distribution-of-energy of electrons after passage through hydrogen, confining

[18] The continuous curves marked A and B are graphs of predicted distribution-curves deduced (for atomic hydrogen) from the assumptions of wave-mechanics by Born. The ordinates of all the curves, experimental and theoretical alike, have been adjusted so that all five intersect at 5°.

their observations however to those which had been deflected only a little or not at all. They found that many suffered an energy-loss of about 12.6 equivalent volts, evidently the same which Harnwell was later to observe. Arnot used Dymond's method in a study of mercury vapor; he detected energy-losses corresponding to excitation and ionization, and traced curves resembling those of Fig. 11. Rudberg, working with nitrogen, and likewise concerning himself only with

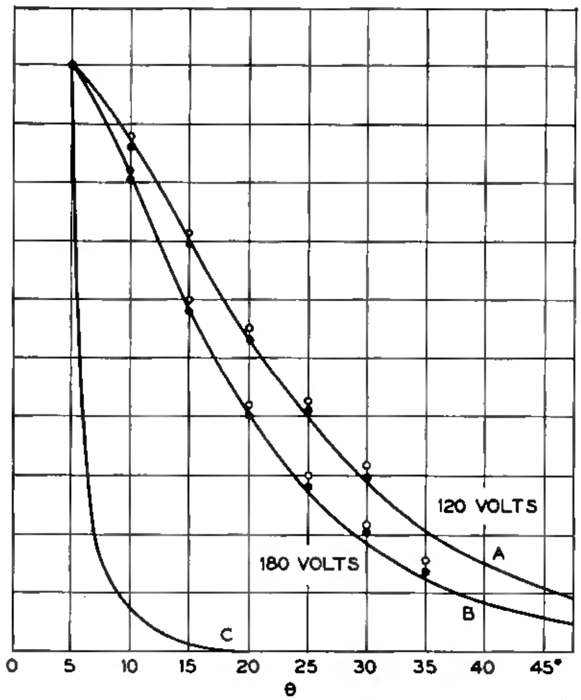


Fig. 11—Distribution-in-angle of electrons scattered from molecules and atoms of hydrogen. (G. P. Harnwell, *Physical Review.*)

electrons which were practically not deflected at all in their collisions, obtained curves with maxima remarkably sharp, indicating several distinct and very precisely determined energy-transfers, the two most prominent amounting to 12.78 and to 9.25 equivalent volts.

Yet another device is that of R. Kollath, which is much simpler than Harnwell's, but functions only for one small range of angles of scattering; it is shown in Fig. 12. The primary electron-beam passes through the slits B_1 and B_2, then onward into the collecting-chamber A; those corpuscles which are scattered at angles between (roughly) 87° and 93° are able to pass between the flat metal rings L of which one sees the traces on the plane of the paper, and go on to the annular collector K; those which are deflected through other angles are caught by the rings. A potential-drop from the rings to the annular collector precludes from reaching this the electrons which in being scattered have lost more than 15 per cent of their initial energy, so that the ratio of the currents to K and A is a measure of the proportion of the

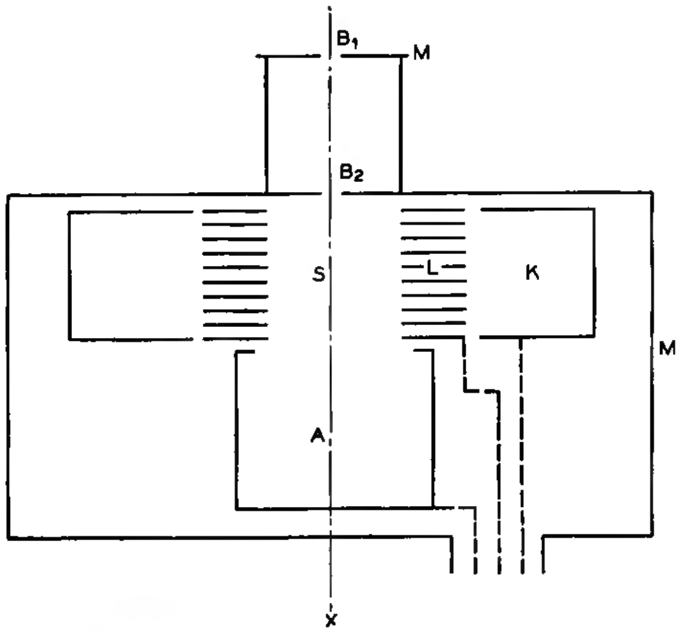


Fig. 12—Kollath's apparatus for determining amount of scattering at angles near 90°. (*Annalen d. Physik.*)

primary electrons, which in traversing a (known) distance through gas of a (known) density are deflected without more than the stated loss of energy through angles within 3° of a right angle. Measuring the ratio at various pressures of gas, Kollath determined the corresponding cross-section for a number of gases. The continuous curves of Fig. 13 are those which he obtained for three of them; the broken curves represent the intercepting cross-section as found by Ramsauer, with ordinates reduced in the ratio 1 : 20. One sees that for each gas the two cross-sections vary in much but not altogether the same way.[19]

I must quote the results of the work of Langmuir and Jones, though

[19] It happens that the ratio 1: 20 is about the ratio which the electrons received by the collector K would bear to all the scattered electrons, if the distribution-in-angle of these were isotropic—which of course is not necessarily so.

the method which they invented is so extremely different from any of those by which the foregoing data were acquired, that in this place I cannot give anything like a full account of it. Briefly: the gas is in a metal cylinder having a filament running along its axis and metal plates

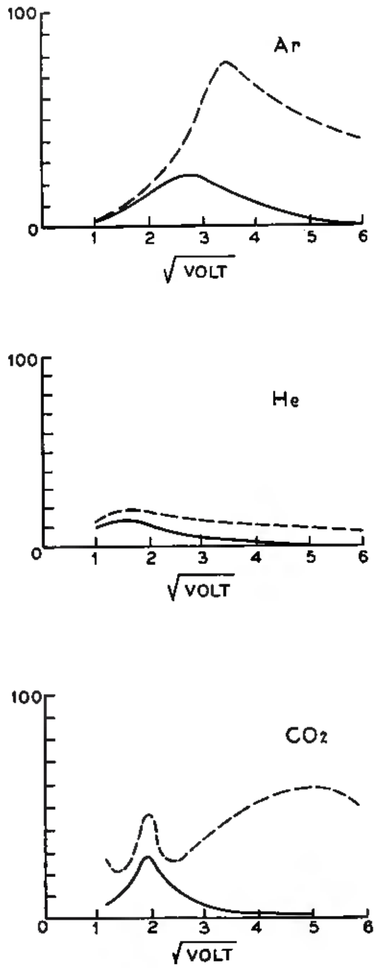


Fig. 13—Broken curves: cross-sections by Ramsauer's method. Continuous curves: cross-sections for scattering at angles near 90°, by Kollath's method (see text). (*Annalen d. Physik.*)

which almost close its ends. One of these end-plates is raised to a potential some fifty or a hundred volts above the filament; and so dense is the electron-current pouring out of this latter, that almost the whole of the gas in the cylinder becomes violently ionized and shining, and assumes the potential of the plate. I say "almost all" of the gas; a narrow cylindrical sheath about the wire remains comparatively dark, and between the wire and the outer frontier of this sheath the entire potential-rise is spread. Having traversed the frontier, the electrons shoot into the luminous zone of the gas with the energy corresponding to the full potential-rise from the filament to the end-plate. It is as though the boundary of the sheath were a grid connected to the plate; the discharge itself creates its own impalpable grid. The method, then, consists in measuring the current into the cylinder over a range of values of the potential thereof, beginning when the cylinder is at the same potential as the filament and the only electrons which can attain it are those which come clear through the luminous zone without deflection or loss of energy, and ending when it is at the same potential as the end-plate and the luminous gas (or at any rate when it is well above the filament) and electrons can reach it despite their collisions en route. By analyzing the shape of the curve, Langmuir and Jones are able to deduce the values of the cross-section for interception for the various gases they tested, and values of several other things as well; but the analysis is intricate. I shall therefore say only that for the gases neon, hydrogen, argon, helium, nitrogen and mercury, and applying to the electrons voltages ranging from 30 to 100, they found for the cross-section values departing by less than ten per cent from the gas-kinetic cross-section σ_0.

Interception of Positive Ions (Atom-nuclei and Charged Atoms)

We consider next the results of experiments like these of the foregoing pages, in which the beam traversing the gas is a stream of positive ions—protons, or atoms of heavier elements lacking each an electron—and the quantity measured is the number of ions disappearing from the beam, or something more or less nearly equivalent. The experiments as yet are few, and mostly not so accurate as those on electron-beams. This is partly because of the difficulty of obtaining steady reliable sources of positive ions—a difficulty which amounts almost to an impossibility, except for alkali-metal ions, and particles issuing from high-voltage discharge-tubes with kinetic energy amounting to thousands of equivalent volts.

The cross-section for interception, σ, is defined in the same way as for electrons—as the ratio $(Q - R)/NQdx$, to return to the notation of equation (1) of this article. But unless the positive ion is a proton, we should certainly not visualize σ as the cross-sectional area of a molecule of the gas; the size of the flying ion itself is involved. The customary procedure is to compute the "radius" R corresponding to the "area" σ by the formula $R = \pi\sigma^2$; then to divide R into the "radius r_0 of the molecule" and the "radius r_i of the ion," according to some more or less plausible guess; then to compute the value of πr_i^2, and call it the cross-section of the ion. But all such procedures are more or less dubious, whereas the quantity σ always retains its meaning as a measure of the likelihood of interception.

There is a particular mode of interception for positive ions, different from any to which electrons are subject. If an ion passes close to an atom, it may steal an electron to neutralize itself. The former atom becomes of course an ion, but in general it does not acquire the velocity of the former ion, while the latter being neutralized is not perceived even if it continues on its course to the collector. Therefore, the net result of the electron-transfer is the vanishing of an ion from the beam. It follows from quantum-mechanics that the cross-section of the atom for this special sort of interception is greater, the more nearly the energy required to ionize it agrees with the energy required to re-ionize the neutral atom into which the former ion is transformed by the electron-transfer. Consequently, the greatest values occur when the ion-stream consists of ionized atoms of the very gas through which it is being sent.[20]

One would wish, other things being equal, to duplicate with positive ions the apparatus already used with electrons. The best example of a duplication is that presented by Ramsauer and O. Beeck, who used the device of Ramsauer's earlier work depicted in Fig. 2, substituting for the metal plate at Z a metal ribbon painted with an amalgam of mercury with any of the five alkali metals; heating the ribbon, they got a stream of the ions Li^+, K^+, Na^+, Rb^+, or Cs^+. The gases which they used were A, He, Ne, H_2, N_2, and O_2, though of these argon was the only one of which they measured the interception for all five kinds of ions. The range of energy-values extended from 1 to 30 equivalent volts, and over it the value of σ in every case diminished steadily with increasing energy. When different ions were driven through the same gas, the value of σ was found to be greater, the greater the atomic number of the ion; when ions of the same kind were tried in different gases, it was found to be greater, the greater the molecular weight of the gas. All this is as one would expect. The value of R (as defined in the paragraph above) is of the same order of magnitude as the sum of the gas-kinetic radius of the molecule or atom of the gas, and that of the ion.

[20] For the theory, and for a bibliography of the experimental work, see H. Kallmann & B. Rosen, *ZS. f. Phys.* **61**, 61–86 (1930). I shall treat the subject more extensively elsewhere.

Dempster experimented with a scheme which may be likened to Ramsauer's with the final slit halfway around the circle from the initial slit, and all the slits and walls between suppressed.[21] Later it was adopted by some of his pupils, and by Kallmann and Rosen. Various mixtures were used by Dempster and his pupils to produce the ions; some were heated, some bombarded by electrons; one of them, when thus bombarded, emitted protons and H_2^+ ions, a very useful property. When one varies the magnetic field and plots against it the current which passes through the final slit and is captured by an electrometer posted behind, one gets a curve with a series of peaks, one for each kind of ion present.

When the deflection-chamber is well evacuated, the peaks are sharp and narrow; when gas is introduced, they become lower and broader, and sometimes they are visibly displaced. In certain cases (K^+ ions in helium, for instance) the height of the peak falls off exponentially with increase of pressure, and one deduces the value of σ by equation (9); it invariably turns out to be smaller than the value one gets by supposing both the ions and the atoms to have their gas-kinetic cross-sections, and decreases as the speed of the ions is increased. The data are thus in qualitive agreement with those of Ramsauer and Beeck, but a thoroughgoing comparison is yet to be made. The broadening of the peaks is a sign that the ions in their encounters with the particles of the gas are suffering small deflections with scarcely any loss of energy; the displacement of the peaks (when it occurs) a sign that they are losing small amounts of energy. A simple reduction in the height of a peak, unattended by broadening or displacement, might well indicate that the electron-transfers far outweigh the other modes of interception, and that the value of σ obtained is the σ for electron-transfer.

The value of σ for protons, Dempster found, is remarkably small. One may visualize these hydrogen nuclei as mere points, as one does an electron, and consider σ as the "cross-section for interception of protons" of molecules of the gas. Sending 900-volt protons through helium, in apparatus of the type above, Dempster observed the peak surviving even when the density of the gas was so great, that if σ had been equal to the gas-kinetic cross-section σ_0, more than half of the ions would have been intercepted in the first one-hundredth of their semi-circular path. It was considerably broadened, as though most of the protons had suffered small deflections; but at a density as much as one-seventh as great, there was hardly any broadening even;

[21] I must not give the impression that Dempster's scheme was a copy of Ramsauer's or Smyth's; it antedated both, having been used for other purposes, *e.g.* the detection of isotopes.

and at the high density, few had been deflected enough to be caught by the walls.[22]

G. P. Thomson developed another way of studying what happens to a stream of protons shooting through a gas. He set a photographic plate athwart the path of the narrow beam, and observed the imprint, which grew broader when a gas was introduced into the path. By measuring the darkening of the plate from point to point across the imprint, he was able to deduce the law of scattering-in-angle of the protons. The most interesting single feature of his data was a likeness to Ramsauer's famous observation with electrons: using either argon or helium, Thomson found that as the speed of the protons was decreased, the broadening of the beam rose to a maximum and then declined. This implies that for either gas the cross-section for interception of protons has a maximum, like that for interception of electrons. Moreover for either gas, the two cross-sections have their maxima at about the same *speed,* though not the same *kinetic energy*—owing to their difference in mass, a proton has some 1840 times as much vis viva as an electron keeping pace with it. Those with which Thomson was working had energy measured in thousands of volts.

Much of the work of Kallmann and Rosen is directed to testing the theorem stated above about the cross-section for electron-transfer. This they succeed in doing on a number of cases, with favorable results. Thus in nitrogen gas, the value of σ is much greater for N_2^+ ions than for N^+ ions; but in oxygen it is the N^+ ion for which the value of σ is greater, for the ionizing potential of the N atom lies closer to that of the O_2 molecule than does that of the N_2 molecule. Likewise Holzer has found that in hydrogen, the cross-section for interception is greater for H_2^+ than for either H^+ or H_3^+. It is not certain in any of these cases that the σ measured refers entirely to interception by electron-transfer, but the evidence nevertheless is strong.

The scattering of exceedingly fast helium nuclei—alpha-particles—is a subdivision of this field,—the best known, probably, of all, since out of it the central feature of our contemporary theory of the atom is derived. The value of σ for this scattering is vanishingly small, by comparison even with the gas-kinetic cross-section. But since the speed of the particles is so great, this result is for once exactly what we should expect.

[22] Dempster phrases the results in terms of the "mean free path," the reciprocal of $N\sigma$ (N standing for the number of atoms of gas per unit volume). The "high density" mentioned in this sentence was such, that the semi-circular path of the protons between the two slits was 108 times as long as the quantity $(N\sigma_0)^{-1}$ previously defined, the "gas-kinetic mean free path." Dempster also observed the extinction of a beam of H_2^+ ions, accompanied by the advent of a beam of slow-moving H^+ ions probably due to the dissociation of the former.

REFERENCES

(*A*) Measurement of Cross-sections for Interception

A. Becker: *Ann. d. Phys.*, **17**, 381–470 (1905); **67**, 428–461 (1922).
W. H. Brattain: *Phys. Rev.* (2), **34**, 474–485 (1929).
R. B. Brode: *Proc. Roy. Soc.*, A, **109**, 397–405 (1925); *Phys. Rev.* (2), **25**, 636–644 (1925); *Phys. Rev.* (2), **34**, 673–678 (1929); see Normand.
E. Brüche: *Ann. d. Phys.*, **81**, 537–572 (1926); **82**, 25–38, 912–946 (1927); **83**, 1065–1128 (1927); **84**, 279–291 (1927); *Ergebnisse d. exakt. Naturwiss.*, **8**, 185–228 (1929).
M. C. Green: *Phys. Rev.* (2) **36**, 239–247 (1930).
T. J. Jones: *Phys. Rev.* (2), **32**, 459–466 (1928).
R. Kollath: *Ann. d. Phys.*, **87**, 259–284 (1928); see Ramsauer.
I. Langmuir, H. A. Jones: *Phys. Rev.* (2), **31**, 357–404 (1928).
P. Lenard: *Wied. Ann.*, **56**, 255–275 (1895); *Ann. d. Phys.*, **12**, 714–744 (1903).
L. R. Maxwell: *Proc. Nat. Acad. Sci.*, **12**, 509–514 (1926).
H. F. Mayer: *Ann. d. Phys.* **64**, 451–480 (1921).
C. E. Normand: *Phys. Rev.* (2) **35**, 1217–1225 (1930).
C. E. Normand, R. B. Brode: *Phys. Rev.* (2), **35**, 1438 (1930).
C. Ramsauer: *Ann. d. Phys.* **64**, 513–540 (1921); **66**, 546–558 (1921); **72**, 345–352 (1923); *Phys. ZS.*, **29**, 823–830 (1928).
C. Ramsauer, R. Kollath: *Ann. d. Phys.* (5), **3**, 535–564 (1929); (5), **4**, 91–108 (1930).

(*B*) Analysis of Distribution-in-Energy of Scattered Electrons

F. L. Arnot: *Proc. Roy. Soc.*, A, **125**, 660–669 (1929).
E. G. Dymond, E. E. Watson: *Proc. Roy. Soc.*, A, **122**, 571–582 (1929).
G. P. Harnwell: *Phys. Rev.* (2), **33**, 559–571 (1929); **34**, 661–672 (1929); **35**, 285, (1930).
H. Jones, R. Whiddington: Phil. Mag. (7), **6**, 889–910 (1928).
J. H. MacMillen: *Phys. Rev.* (2), **36**, 1034–1043 (1930).
E. Rudberg: *Nature*, **125**, 165–166 (August 2, 1930).

(*C*) Interception of Protons and Other Positive Ions

W. Aich: *ZS. f. Phys.* **9**, 372–378 (1922).
A. J. Dempster: *Phil. Mag.* (7), **3**, 115–127 (1927).
F. M. Durbin: *Phys. Rev.* (2). **30**, 844–847 (1927).
R. E. Holzer: *Phys. Rev.* (2) **36**, 1204–1211 (1930).
H. Kallmann, B. Rosen: *ZS. f. Phys.* **61**, 61–86, 332–337 (1930); **64**, 806–816 (1930); *Naturwiss.* **18**, 448–452 (1930).
R. B. Kennard: *Phys. Rev.* (2), **31**, 423–430 (1928).
C. Ramsauer, O. Beeck: *Ann. d. Phys.* (4), **87**, 1–30 (1928).
G. P. Thomson: *Phil. Mag.* (7), **1**, 961–977 (1926); **2**, 1076–1084 (1926).

A Study of Telephone Line Insulators *

By L. T. WILSON

This paper discusses the major factors contributing to the (total) leakage conductance of telephone line insulators, especially at carrier frequencies up to 50,000 cycles. The influence of both the design and material of the insulators and pins on each factor is discussed and illustrated by test data.

The electrical performance of three different designs is analyzed to illustrate, in a general way, the relative importance of the several factors.

To the layman, the long strings of large power insulators suspended from tall steel towers naturally present a more imposing picture than do the small telephone insulators mounted on the crossarms of relatively short wooden poles.

To the engineer, however, these little insulators present problems quite as stimulating and interesting in the telephone field as do the large insulators in the power field.

It is the purpose of this paper first to discuss briefly how some of these interesting problems arose and then to cover in more detail a study of the major phenomena involved in insulator leakage and finally to show how the knowledge gained from that study has helped bring about improved telephone insulators, some of which will be described.

Origin of Problem

For many years the requirements of telephone insulators were relatively easy to meet because the frequency of the currents transmitted did not exceed about 3 kc. and because the leakage of insulators is generally low at such frequencies.

Therefore, the familiar glass insulators such as are shown in Figs. 1 and 2 sufficed, the former design (D. P. type) being employed on the longer circuits and the latter (toll type) on the shorter ones. Indeed they still suffice very generally, especially where only currents of voice frequencies or less are transmitted.

The advent of carrier systems employing higher frequencies ranging from about 3 to 30 kc. changed the insulator requirements substantially. At first these systems were few in number and relatively short in length and the insulator problem accordingly less important.

* *Presented at the Summer Convention of the A. I. E. E., Toronto, Ont., Canada, June 23–27, 1930.*

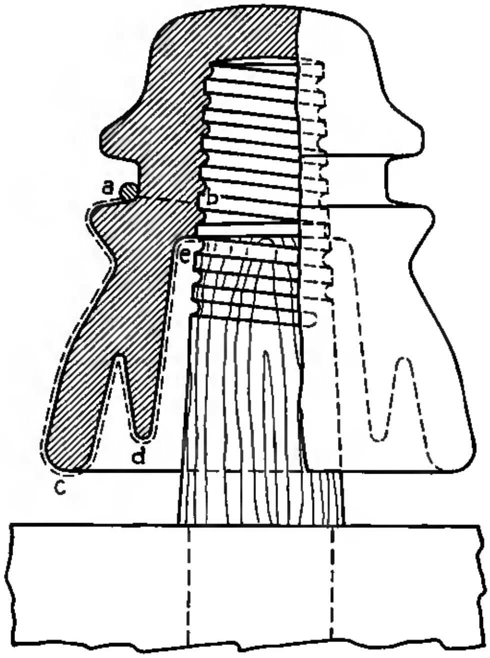


Fig. 1—Standard D. P. insulator and standard wood pin.

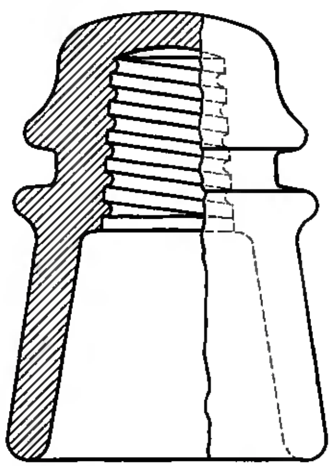


Fig. 2—Standard toll insulator.

The rapid growth of carrier systems during the past decade,[1] the longer circuits to which they have been applied, and improved standards of long-distance transmission have all been factors in increasing the requirements imposed upon the insulator and in correspondingly augmenting the importance of the problem.

It may now be remarked that the problem has been mainly one of securing economical insulators giving improved performance at these higher frequencies. In addition the low frequency performance of new designs had to be maintained substantially as good as that of the old designs.

Leakage Phenomena

This study has been confined almost entirely to the pin type of insulator.

When an alternating potential exists between a pair of wires at the point where those wires are supported by insulators, a current flows from the one wire to the other. This current may be resolved into two components, one in phase with the potential and one in phase quadrature leading the potential.

This in-phase component which, of course, represents an energy loss is the one of chief interest here and in using the word leakage we refer to this component, or more accurately to its equivalent conductance.

Of course, both components in flowing through the resistance of the line conductors produce energy losses but these are so small as to be omitted here. Except these, all other energy losses which occur due to the presence of insulators whether they actually occur in the insulators proper or elsewhere will be charged to insulator leakage.

It is convenient to divide insulator leakage into several sources. This division is an arbitrary one, because some of the sources are not independent of each other and are, as will be seen later, difficult to separate experimentally.

The division follows:

A.-c. and d.-c.	A.	Leakage directly through insulator material to pin.
	B.	Leakage directly over insulator surfaces from line conductor to pin.
A.-c. only.	C.	Dielectric absorption in insulator material.
	D.	Dielectric absorption in pins.
	E.	Displacement current losses in crossarms and pins.
	F.	Losses due to unbalanced displacement currents in external resistances such as those of crossarms, poles, etc.
	G.	Displacement currents flowing over insulator surfaces through high resistance.

It should be noted that while all the items play a part in a.-c. leakage only the first two enter in the d.-c. case.

[1] "Carrier Systems on Long Distance Telephone Lines," H. A. Affel, C. S. Demarest, and C. W. Green, *A. I. E. E. Trans.*, Vol. 47, 1928, pp. 1360–1386.

These items[2] will now be discussed individually and will be treated generally as they exist under wet rather than dry weather conditions. In this connection most of the experimental evidence that will appear was obtained from tests on insulator test lines located near Phoenixville, Pa. The measuring equipment employed there has already been described in another paper[3] and for lack of space will not be further discussed here although certain improvements have been made since that paper was prepared.

Item A—Direct Leakage To Pin

This item refers to that leakage through the insulator material which is directly due to the conductivity of the material, for example via path *a b*, Fig. 1. Item *A* has been listed chiefly for completeness because it is fairly well known that for most materials commonly employed, this source of leakage is very small. In the present study it has been found to be negligible in magnitude.

Item *B*, therefore, becomes the controlling source of leakage for direct current and accordingly deserves consideration. However, it will be given even more consideration because its magnitude, being so closely dependent on the shape, size, location, and condition of the insulator surfaces throws light on those factors which will later be shown to play a very large part in leakage for alternating currents.

Item B—Direct Surface Leakage

1. General Characteristics. Item *B* refers to that leakage over the insulator surfaces which is directly due to the conductivity of those surfaces, for example, via path *a c d e*, Fig. 1.

The outstanding characteristic of this type of leakage is the enormous changes in magnitude it exhibits under changing weather conditions.

For example, Fig. 4 shows the results of measurements made on a telephone line in Texas equipped with insulators of the type shown in Fig. 3. This test, made in clear weather, shows the large change in leakage produced simply by the condensation of moisture on the insulator surfaces. The peak value may be seen to be about 2500 times the smallest value measured.[4] Had it rained along the entire length of line the peak might readily have been 10 times greater and the correspond-

[2] The existence of several of these factors appears to have been well appreciated by Mr. R. D. Mershon as early as 1908, although Mr. Mershon's measurements were made at frequencies below 100 cycles where many of the items are extremely small in magnitude. See "High Voltage Measurements at Niagara," by R. D. Mershon and Discussion, *A. I. E. E. Trans.*, Vol. 27, 1908, pp. 845–929.

[3] "Methods of Measuring the Insulation of Telephone Lines at High Frequencies," E. I. Green, *A. I. E. E. Trans.*, Vol. 46, 1927, pp. 514–519.

[4] Such tests indicate the negligible effect of Item *A*.

ing ratio would have been 25,000. These ratios are higher than would be commonly found because the line was quite new but they serve to indicate the wide range in the magnitude of this type of leakage.

Fortunately, direct surface leakage even at its higher values in a sufficiently small part of the total leakage, especially at carrier frequencies, to make its wide variations substantially less serious than the foregoing ratios have indicated.

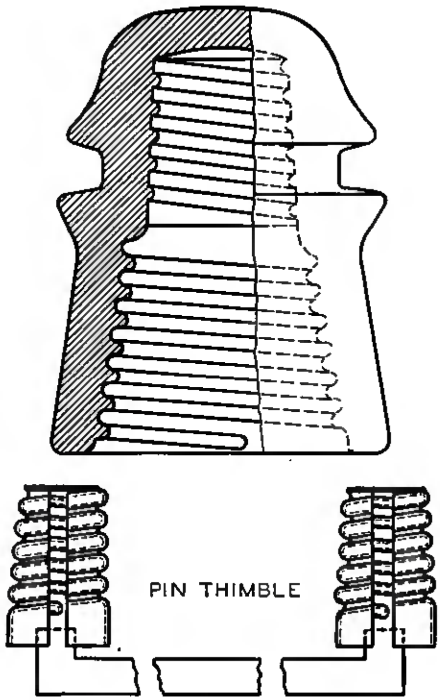


Fig. 3—Standard C. W. insulator and pin thimble.

2. *Influence of Insulator Design.* To make this leakage conductance small it will be apparent from elementary considerations alone that the length of the path should be as great and its cross-section as small as possible. The latter depends on both the thickness and width of the moisture film, the width, in turn, depending upon the diameter of the surfaces.

The length of path may be increased by simply lengthening the insulator. The cross-section of the film may be made small by decreasing the diameter of the surfaces. Both methods are employed in the experimental design shown in Fig. 5 which will be recognized as extreme, at least with respect to mechanical strength. This design under test gives excellent performance for direct surface leakage, having under various weather conditions from one to 25 per cent as much leakage as the more reasonable design shown in Fig. 6.

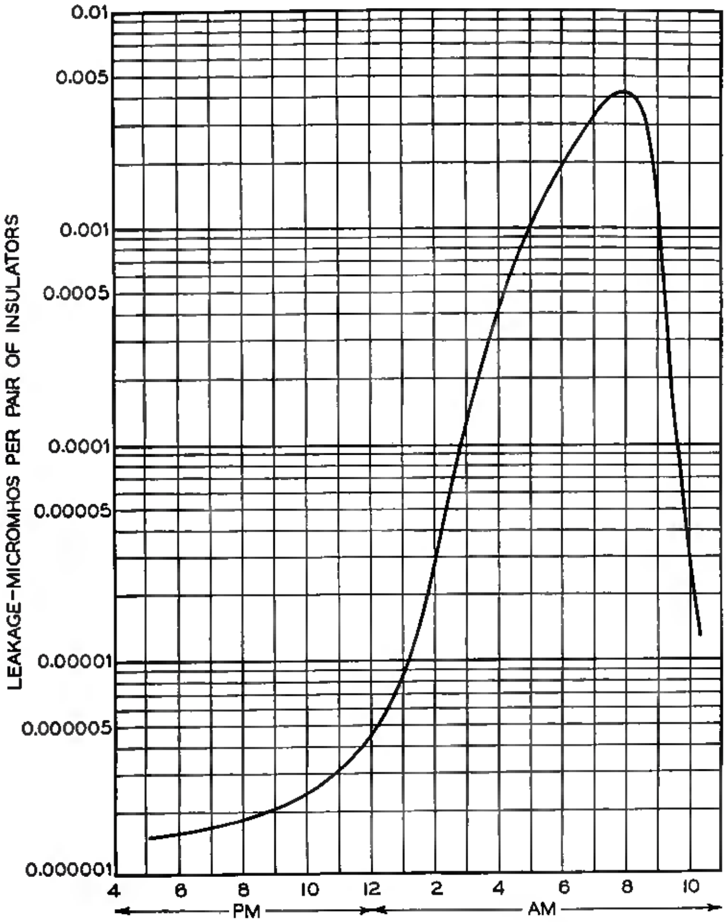


Fig. 4—Illustrates large variation of d.-c. surface leakage in dry weather.

Another, but somewhat less effective, way of increasing the length of path is the use of one or more petticoats. This method has a mechanical advantage in keeping the overall insulator length small but the effectiveness of the longer path is partly lost due to the necessary increase in diameter. An example of this method is the design shown in Fig. 1 which gives about half as much leakage as that of Fig. 2.

Still another way of increasing the length of path is the use of corrugations (Fig. 6, for example). The data at hand while not conclusive, indicate that corrugations are of questionable value.

The third dimension of the conducting path, namely its thickness, can be controlled to some extent by insulator design. The thickness of the water film is determined by the rate of rainfall and by a balance

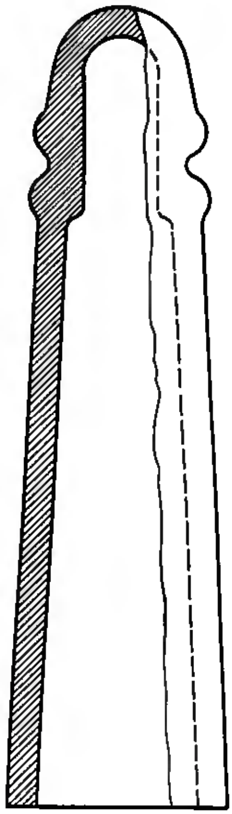


Fig. 5—Experimental design with long skirt.

between the forces tending to make the film adhere to the surfaces and the force of gravity tending to pull the water away.

To reduce the amount of water intercepted by the insulator for a given rate of rainfall, it is again advantageous to make the insulator diameter as small as possible.

To facilitate the running off of the water intercepted, the surfaces in the conducting path should be nearly vertical and smooth. These remarks apply mainly to the outside surfaces.

As to the inside surfaces the situation is somewhat different. These are protected from direct rainfall and become wet by splashing, condensation, creepage, and by moisture carried by air currents.

To prevent splashing from the crossarm it might appear desirable to provide the insulator with a wide flaring kind of shed but experience has shown that such sheds may be quite ineffective or even detrimental. It will be convenient to discuss the action of one type of shed later.

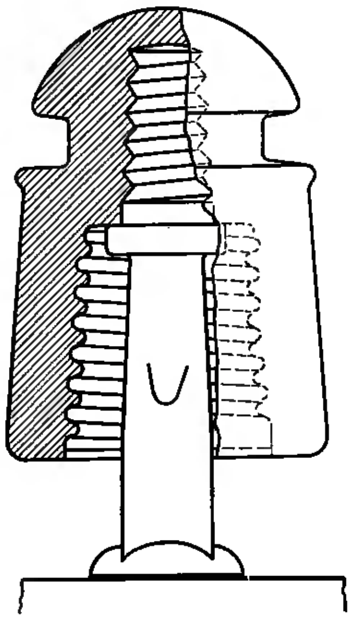


Fig. 6—Experimental design with short corrugated skirt (C. P.) mounted on standard steel pin.

It has been found preferable to employ an insulator of small diameter. The effects of the splashing can be reduced by placing the insulator higher from the crossarm and by restricting the area of the opening between the pin and the insulator where the rising drops of water must enter.

These drops may readily rise a foot or more as casual observations show. To raise an insulator completely out of range obviously would lead to a cumbersome and mechanically unsatisfactory construction. However, some advantage may be had in elevating the insulator.

For example, consider Fig. 7 which shows the insulator of Fig. 6 mounted on a long pin. This pin, which was designed for another purpose, has an enlarged section which restricts the area between the pin and the insulator. The improved performance is, therefore, the result of two effects: one, that of elevating the insulator; the other, that of restricting the area between the pin and the insulator.

An attempt was made to separate these effects by providing a third set-up in which the insulator is mounted on the pin of Fig. 6, this short pin being supplied with a rubber washer of the proper size and location to simulate the enlarged section of the long pin.

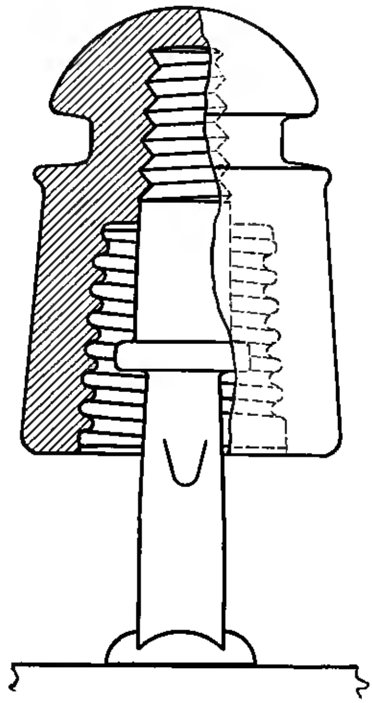


Fig. 7—C. P. design on long steel pin with baffle.

The relative performance of these three arrangements is given in Fig. 8 where *S* stands for short pin; *SB*, short pin with baffle; and *LB* long pin with baffle. These results were obtained during a hard shower within a month after the installation of the rubber washers and give some idea of the effectiveness of the two expedients.

In discussing the influence of insulator design on this type of leakage the question naturally arises as to the part played by the pin on which the insulator is mounted.

Wood pins of the type commonly employed on telephone lines are found to contribute somewhat to the total insulation, depending on the dryness of the wood and on the efficiency of the insulator itself.

The effect of the pin may be measured by comparing insulators

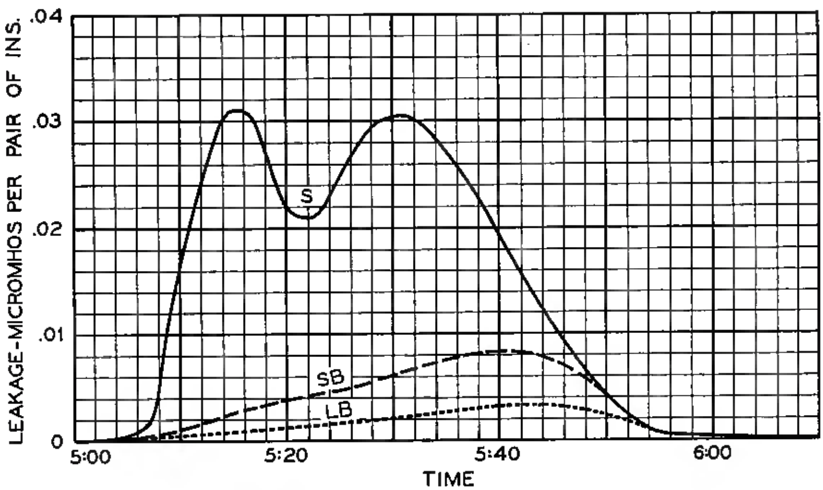


Fig. 8—Illustrates effect of pin length and baffle on d.-c. surface leakage.

mounted on wood pins in the standard manner with similar ones mounted on wood pins which have been covered with a thin metal foil,

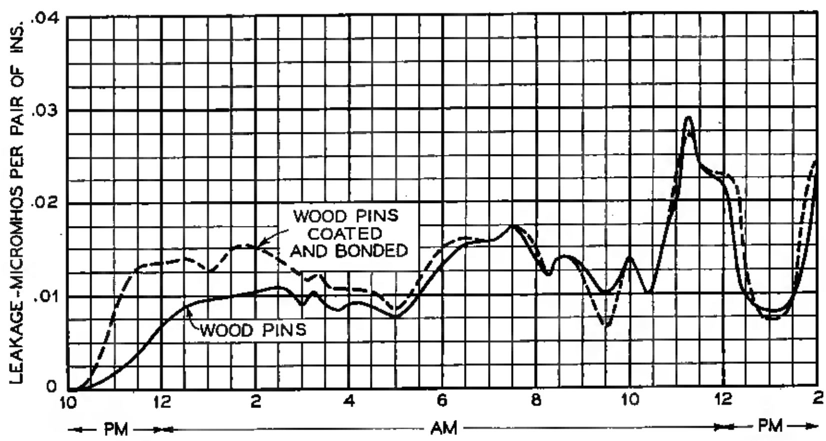


Fig. 9—Illustrates effect of coating and bonding of wood pins on d.-c. surface leakage.

the coatings of a pair of pins being electrically joined. Thus, both the pins and the crossarm between them are short circuited.

Fig. 9 shows the results of such a test on the type of insulator shown in Fig. 1. At the start of the rain and during the first few hours the wood appears to reduce the leakage considerably but as the rain continues and the pin takes up moisture its effect seems to decrease to nil.

While similar tests on this and other types of insulators give corresponding results, still other tests indicate that at times, the wood continues to help out for many hours. One such test covering a period of about 23 hours showed, for this same type of insulator (Fig. 1), that the wood pin reduced the leakage by 30 per cent on the average.

On the other hand, similar tests for the design of insulator shown in Fig. 10, very frequently show the pin to have negligible effect.

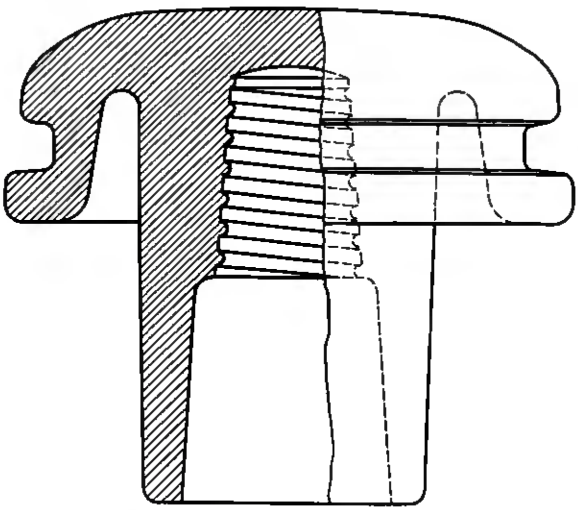


Fig. 10—Experimental design (C. D.)

While the foregoing tests have shown an advantage in favor of the wood pin, that conclusion holds only for the conditions of the test; namely, both wood and metal pins having the same diameter. When an insulator is specifically designed for a metal pin both the pin and insulator diameters can usually be made small because of the greater mechanical strength of the metal pin. The advantages of the design of small diameter, as previously discussed, may readily offset the slight disadvantage of the metal pin and the net result may be an actual gain.

3. *Influence of Insulator Material.* While the insulators are new, the molecular attraction between the insulator material and water plays a very important part; so important, in fact, that tests on such new insulators frequently give an unreliable basis on which to make decisions. On new glass insulators and particularly on those of

borosilicate compositions, the rain water does not seem to wet the surfaces but rather stands in individual separate drops. The phenomenon is so marked that the surfaces have the appearance of having been oiled or waxed.

Under such conditions the conducting path is broken up and is discontinuous. Naturally its conductivity is very small. It would be fine if this property of the material could be preserved. Unfortunately, along with exposure to weathering there comes an increasing tendency for the drops of water to spread out and unite to form a continuous path. Apparently, this action results from the collection of impurities on the surfaces, for such molecular phenomena are well known to be very sensitive to any contamination.

The direct surface leakage has been observed generally to increase 10 or more times after only a few weeks' exposure. In one particular case where it rained the same day that the new insulators were installed, their surface leakage was observed to be less than one per cent of that measured on insulators of the same shape and material which had been aged for about two years.

This study of insulators has shown that while changes in design may affect changes in direct surface leakage by a factor of, say, 10, changes in kind of material affect this leakage by a factor of two or less after long exposure. In general, then, the material may be said to be less important than the design. Purely general considerations indicate this conclusion in view of the surface nature of the phenomenon. Consider an insulator of a given design exposed to the elements for several years. More and more foreign matter will collect on the surfaces as time goes on. It is obvious that as this process continues, the insulator material is becoming less and less important and it is conceivable, at least, that given time enough, this surface of foreign matter would determine the insulator performance irrespective of the material beneath this surface coating. At this point, design is paramount.

However, the material might be expected to influence the aging process in some way. From this viewpoint smoothness and hardness of surface, together with chemical stability, appear to be desirable qualities.

This study has been confined chiefly to various kinds of glass. In general, these have all had smooth surfaces, but they have varied in hardness and chemical stability. For example, borosilicate glasses are said to be generally harder and chemically more stable than the common alkali group.

The relative performance of one sample from each of these respective groups is given in Fig. 11. Both were molded to the design of Fig. 12.

This particular test which was made after more than a year of exposure shows only a small difference between the two glasses. A previous test (exposure of nine months) showed quite the same results while

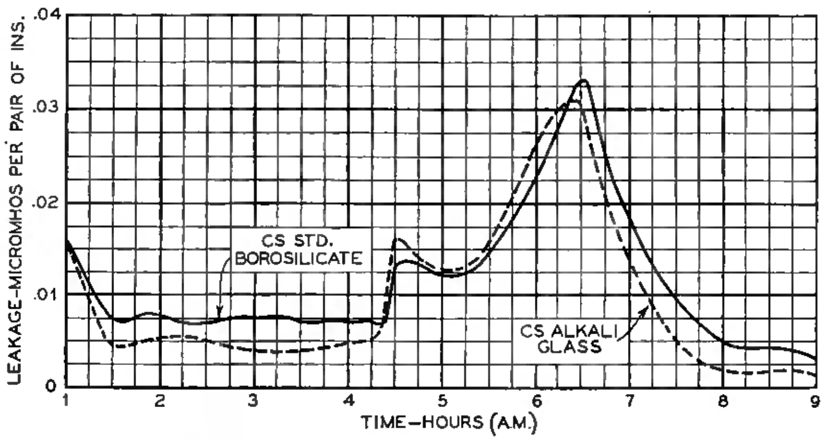


Fig. 11—Illustrates effect of insulator material on d.-c. surface leakage.

still earlier tests had indicated the borosilicate as quite superior to the alkali glass.

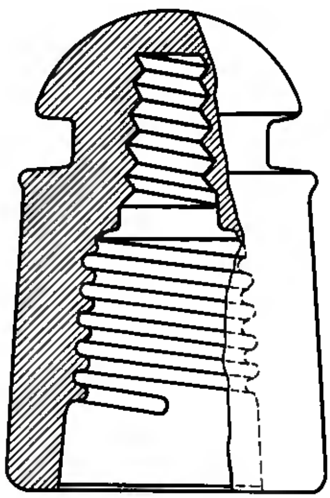


Fig. 12—Standard C. S. design for use with standard steel pin.

While the foregoing tests illustrate the point made above as to the relative importance of design *vs.* material they cannot be considered as conclusive evidence on the relative merits of these two glasses for direct surface leakage.

Probably more important than either hardness or chemical stability, which directly affect surface leakage, is the transparency of the material which affects leakage only indirectly, through the medium of insects, apparently small spiders. These spiders build nests on the inner surfaces and seem to prefer dark or dimly lighted spaces for their homes. Therefore, the more transparent the insulator, the less attractive home it makes; and of opaque materials, probably white ones are accordingly less attractive than dark colored ones.

Other factors which enter into the spiders' choice appear to be those of space and of protection from the elements. These latter factors are functions of insulator design rather than material.

There has been little opportunity to study these factors on the insulator test lines, due to the lack of spiders. Only one specific case has been found where their effect was marked. This was a case where small borosilicate insulators were given an opaque metal coating. After this coating had been on several months the direct surface leakage increased to several times the value for the similar uncoated ones. An investigation showed spider nests under many of the coated insulators.

On the other hand, several types of larger insulators have shown no such effect when coated, although some of these have been exposed for several years. These results are not conclusive. They merely indicate that design, as well as transparency of material, is a factor.

An experience of the telephone plant in the use of opaque insulator material (porcelain) showed a serious reduction of efficiency after a few years of exposure, apparently explained by the action of insects.

4. *Specific Conductivity of Film.* The specific conductivity of the rain itself before it reaches the insulators is determined by the nature and amount of impurities collected in its fall. Both the kind and amount of impurities must vary greatly in different localities, for example, industrial centers as compared with open country. Then again, the amount in any given locality must vary throughout any storm on account of the cleansing action of the rain on the atmosphere.

On reaching the insulators the rain will suffer a further increase in conductivity depending on the impurities it finds there. Smooth vertical surfaces should be advantageous in reducing the collection of dust.

After a prolonged dry period in which the surfaces have become dust coated, the conductivity may be quite high at the start of rain. As the rain continues a certain amount of cleansing action occurs on the unexposed surfaces, depending on the splashing. A decrease in leakage corresponding to this cleansing action has occasionally been observed.

Thus, in a locality where rain is infrequent but where dew, for example, might wet the insulators, the leakage might conceivably be quite high.

This phenomenon can be produced artificially and emphasized by placing roofs over the insulators, thus preventing any direct cleansing action of the rain.

Roofs of sheet metal six inches in height, five inches in width and twelve inches in length were placed over insulators of the design in Fig. 1. The ends were left open to permit passage of line wire.

While still new and clean these protected insulators showed less leakage than the unprotected ones of the same design. However, after a

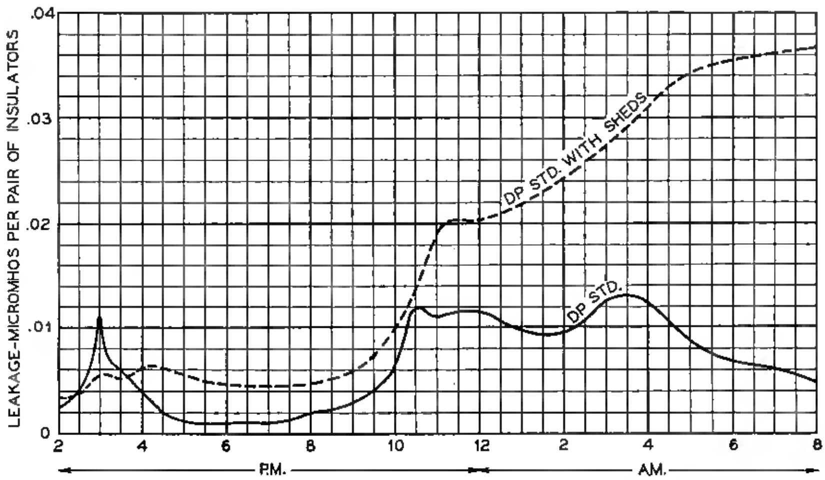


Fig. 13—Illustrates effect of sheds on d.-c. surface leakage.

few months the protected insulators became covered with a thick layer of dust blown under the roofs by wind. Then when it rained, enough moisture reached the surfaces to wet the dust, but not enough to have much cleansing action. The surface leakage then becomes very high, as will be apparent from Fig. 13.

Item C—Dielectric Absorption in Insulator Material

1. General Characteristics. When a dielectric is subjected to a varying potential field a certain amount of the electrical energy is dissipated in the material in the form of heat, depending on the nature of both the material and the field. This phenomenon is commonly called dielectric absorption.

Such a field exists between the line and tie wires on one side and the pin on the other. The insulator material lies more or less in this field and, therefore, dielectric absorption is naturally to be expected. While it is convenient to refer to a single insulator it should be remembered that two insulators are in series between wires at any one point.

The chief characteristic of the leakage resulting from this phenomenon is its variation with frequency. Approximately, it increases directly with the frequency. In the voice range its magnitude is generally negligible, but may become appreciable in the carrier range, particularly at the upper end.

Item C, like item B, increases in wet weather, but to a far less degree. This increase, which is brought about by the enlargement of the field

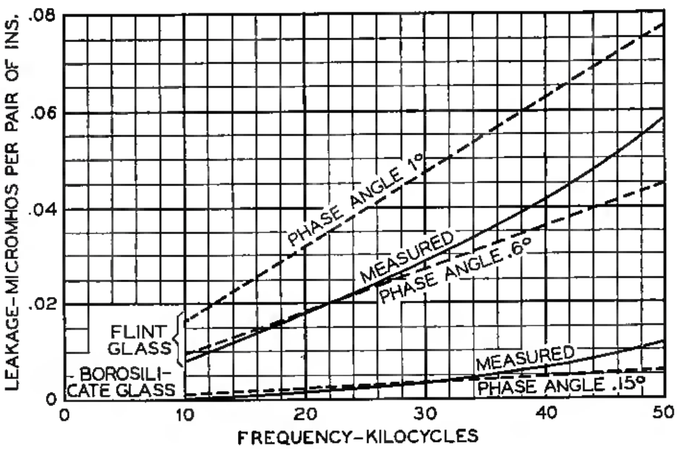


Fig. 14—Variation of (C) with frequency for standard D. P. design.

caused by the wetting of the insulator surfaces, cannot be directly separated from the increases in the several other sources. Indirectly, however, a general idea of the magnitude of item C in wet weather can be obtained in the following two ways.

Both methods are based on the observation that the increase in capacitance between the wire and pin produced by metal-coating the outside of the insulator is invariably greater than the corresponding increase caused by wetting the uncoated insulator. For example, the increase due to coating the insulator of Fig. 3 exceeded the maximum increase observed for any rain by at least 40 per cent. More commonly this figure would be 100 per cent both for this and other shapes.

Therefore, if such metal-coated insulators (mounted on bonded metal pins for reasons appearing in the discussion of items D and E) be measured in dry weather the leakage conductance so found is likely to exceed by a substantial amount the wet-weather value of item *C*.

Another procedure consists in calculating the leakage of the metal-coated insulator. This method requires a knowledge of the wire-to-pin capacitance and the phase angle of the insulator material.

Fig. 14 shows the measured value for the insulator of Fig. 1 when molded from a clear flint glass. The phase angle of this particular glass is not known. However, the wire-to-pin capacitance is known. Using this known value, the leakage has been calculated for two measured values of phase angle for flint glasses, between which the value of this particular glass probably lies.

Similarly, the measured and calculated leakage are shown for this same design of insulator when molded from a borosilicate glass of known phase angle.

Neither method is capable of high accuracy, but they need not be for our present purpose of determining the order of magnitude of item *C*.

Consider item *C* at 50 kc. where its relative importance is greatest. At this frequency the metal-coated flint-glass insulator gave a measured leakage which is only about 10 per cent of the total wet-weather leakage of such an insulator as commonly used uncoated. So it appears that even at this high frequency, item *C* is less than 10 per cent of the total leakage of this particular design and material. This particular sample of flint-glass is not the best of the common alkali glasses nor is it the worst; so this round number of 10 per cent is a fair average value to use for alkali glasses.

The corresponding value for the borosilicate glass of which the sample was representative is about 4 per cent.

2. *Influence of Design.* The absolute magnitude of item *C* is influenced somewhat by design because of the effect of shape on capacitance. An idea of the magnitude of this effect can be obtained by comparing Fig. 14 with Fig. 15. Both designs were cast from the same batch of glass so that material plays a small part in comparing the two designs.

In general, for a given size of pin, the capacitance is decreased by enlarging the insulator diameter, particularly at the wire groove. Similarly, for a given outside diameter, the capacitance is decreased by making the pin diameter less. In general, too, the shorter the insulator, the less the capacitance.

Through the ordinary range of shapes the absolute magnitude of item *C* will not vary more than about three to one, and expressed as a percentage of the total wet-weather leakage, it is doubtful if item *C*, in the worst combination of poor design and poor material studied, could reach as high as 20 per cent.

The foregoing remarks apply to insulators on metal pins. When wood pins are used, the capacitance is less and C is correspondingly less. If the pin is dry, C is extremely small; if the pin is wet, C is still considerably less than it is for the same insulator on a metal pin.

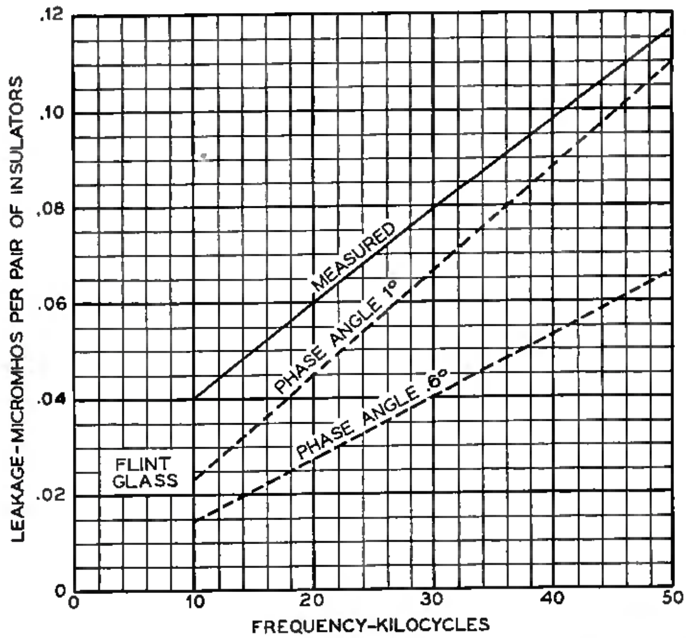


Fig. 15—Variation of (C) with frequency for standard toll design.

3. Influence of Material. The range of the absolute value of C for various kinds of glass is much greater than it is for various designs. Phase angle of the material is frequently used as a criterion but Hoch[5] has shown that for insulator purposes, the product of phase angle and dielectric constant is a better criterion. Using the latter, a range of more than 20 to 1 is found for the glasses studied.

From the standpoint of item C alone, on account of its small relative magnitude there is little justification in going to high grade glasses, especially as these are more costly, unless other sources are first reduced sufficiently to make C really important.

[5] See "Power Losses in Insulating Materials," E. T. Hoch, *Bell System Technical Journal*, November 1922, pp. 110–116.

Item D—Dielectric Absorption in Pins

1. General Characteristics. Item *D* applies only to pins of dielectric material (usually wood) or to metal pins with cobs of dielectric material.

Like *C*, *D* is roughly proportional to the frequency and its importance is greatest at the upper end of the carrier range of frequencies.

Again, like *C*, *D* increases in wet weather due to the accompanying increase in wire-to-pin capacitance. It also increases because of absorption of moisture by the pin.

Fortunately, due to the small value of *C*, particularly when dielectric pins are used, a rough measure of *D* can be obtained by again making use of metal-coated insulators.

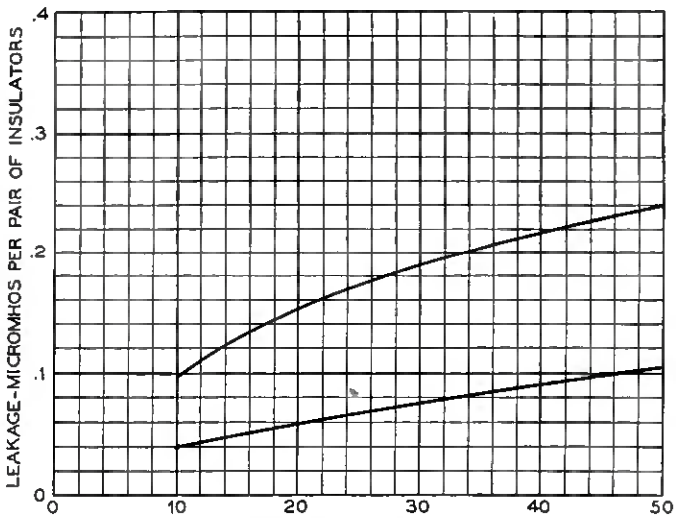


Fig. 16—Variation of (D) with frequency for standard D. P. design.

Fig. 16 shows two measurements of the leakage in dry weather of metal-coated insulators on wood pins. These insulators are similar in shape to that of Fig. 1, and were molded of a borosilicate glass.

Due to the low capacitance when wood pins are used and to the high quality of glass, item *C* can be neglected here.

Item *E*, which is to be discussed in the next section, enters here to some extent, but the cross-section of the crossarm is so much greater than that of the pins that this item, while it cannot be neglected, is probably small.

The measured leakage then closely represents that due to dielectric absorption in the pins under the conditions of the test, namely, metal-coated borosilicate insulators.

To the extent that metal coating gives greater capacitance than does wetting the insulator the value of D is magnified by these measurements. On the other hand, to the extent that the capacitance is reduced by the high quality of glass, the measurements give too low a value of D for alkali glasses.

As these effects approximately balance each other, the measured values are fairly representative of D for the alkali group.

The difference in the two values measured at different times is mainly due to the moisture content of the pins, the higher value corresponding to the higher moisture content.

Consider the insulator of Fig. 1 made of alkali glass and mounted on wood pins. Item C will be small, very small when the pin is dry; and while appreciably larger when the pin is wet, it is still considerably smaller than its value for a metal pin of the same dimensions. D, on the other hand, is large, even when the pin is dry; so in this example, D may be said to be considerably more important a factor than C.

In the case of this same design molded from a borosilicate glass, D becomes relatively even more important a factor.

2. *Influence of Design.* The general remarks which applied to C as to insulator diameter and length and as to pin diameter apply to D.

3. *Influence of Insulator Material.* For a given design, the lower the dielectric constant of the insulator material, the lower D will be.

4. *Influence of Pin Material.* Some study of this subject has been made, but the results are not sufficiently conclusive to report.

Item E—Displacement Current Losses in Crossarms

1. *General Characteristics.* This item will first be considered as it applies to insulators on metal pins. The path of the displacement current is as follows: the current passes from one line wire through the capacitance of the one insulator to its pin, thence through the crossarm to the other pin and finally through the capacitance of the other insulator to the other wire. The capacitances of the two insulators are thus in series. The adjective "displacement" has been applied to the current because its magnitude is mainly determined by the insulator capacitance.

The losses produced in the crossarm by this current obviously depend on the electrical equivalent of the crossarm and on the insulator capacitance. If the former were a pure resistance with a magnitude small compared with the reactance of the series capacitances, then E would increase approximately as the square of the frequency. Experimentally, E is generally found to increase at a rate lying between the first and second power of the frequency over the range studied.

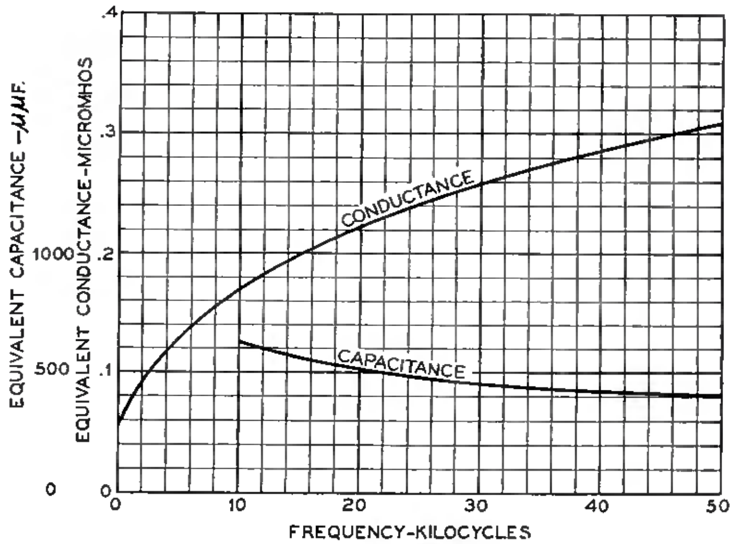


Fig. 17—Variation of equivalent parallel capacitance and conductance with frequency for 25 crossarms in parallel dry weather measurement.

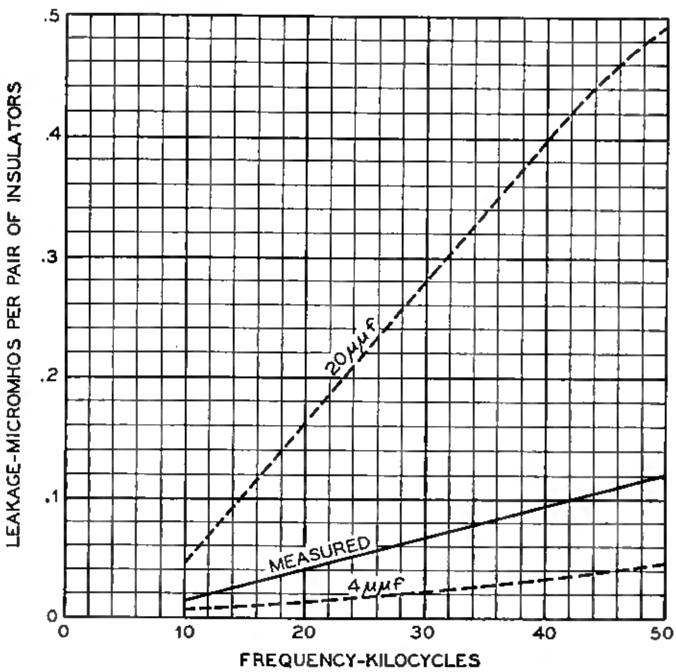


Fig. 18—Variation of (E) with frequency in dry weather.

A knowledge of the electrical equivalent of the crossarm, together with the insulator capcitance, permits E to be calculated.

The constants of the crossarm between steel pins six inches apart, as measured on a dry day, are given in Fig. 17. From this measurement, the magnitude of E has been calculated for individual insulator capacitances of 4 and 20 m. m. f., and the results are shown in Fig. 18. The capacitances of most of the insulators studied lie within these two limits.

Fig. 18 also shows the value of E as measured for the insulators of Fig. 19. This measured value of E was obtained as follows: The leakage of the insulators was measured first with the crossarms short-circuited by tying a wire from each pin to the other. Then the wires were removed and the measurement repeated. The difference between these measurements is E, if other sources are assumed to have remained unchanged. This assumption appears justified from the following test of its validity. The measurement of E, together with the measured constants of the crossarm, enables the insulator capacitance to be calculated. Such a calculation has been carried out and the value of capacitance so obtained checks very closely the value obtained by direct measurement.

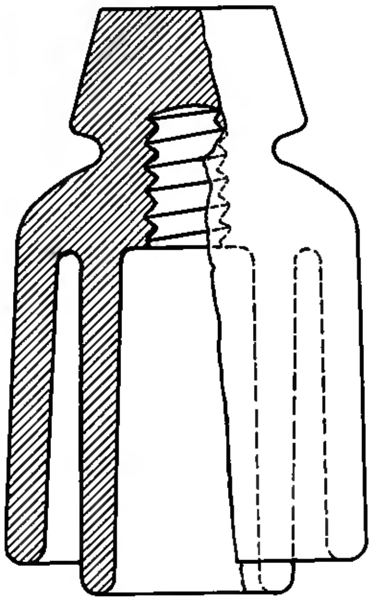


Fig. 19—British post office design.

The curves of Fig. 18 show that E may be large in dry weather. In fact, it is generally several times the magnitude of item C in dry weather and at the upper frequency range. Therefore, E is the controlling factor in dry weather.

As to the effect of weather conditions on E, the increase in insulator capacitance brought by wet weather tends to greatly increase the losses. On the other hand, the large decrease in the crossarm impedance resulting from rain tends to decrease the losses. As a net result of these opposing effects, E is generally less in wet than in dry weather. In fact, after several hours of rain, E has been found to be almost negligible.

The upper curve in Fig. 20 shows E for the insulator of Fig. 1 foil-coated and mounted on steel pins with composition cobs. This curve

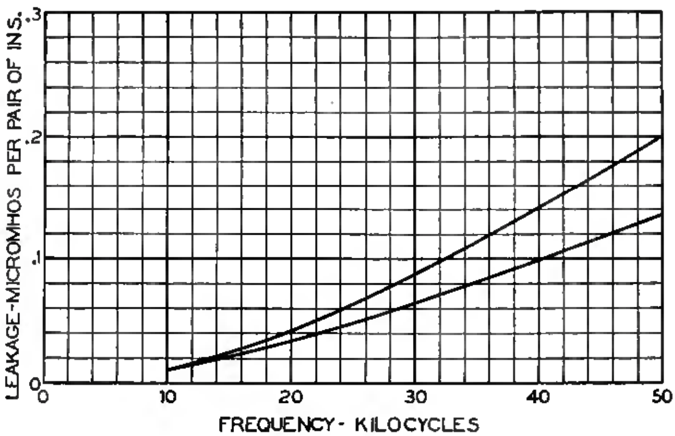


Fig. 20—Measured leakage due to (E).

was obtained in dry weather. The lower curve of Fig. 20 shows the magnitude of E for the same insulators after 1½ hours of rain.

The equivalent conductance of the crossarms varies over a wide range depending on the weather. For instance, values of d.-c. conductance 40 times as great as that recorded on Fig. 17 have been observed and the data indicated that still higher values are probable. The smallest observed value was one tenth of that recorded in Fig. 17.

The experimental determination of the magnitude of E for insulators on wood pins is a difficult problem, because E cannot be readily separated from D.

The importance of insulator capacitance on E has already been established, so a knowledge of this factor for insulators on wood pins enables an estimate of E to be made.

For example, the insulator of Fig. 1 on wood pins gives a measured capacitance in dry weather of about 4 m. m. f. E for this value is given by the bottom curve of Fig. 18. In this particular example, E is quite small in dry weather. The general decrease in E accompanying wet weather indicates that this item does not contribute much of the leakage of insulators on wood pins, at least after the crossarm is very wet.

2. *Influence of Insulator Design.* The wire-to-pin capacitance varies perhaps three to one for the designs covered in this study. The corresponding range in E is obviously great. Therefore, insulator design has an important influence on the magnitude of E.

3. *Influence of Insulator Material.* The materials studied have not shown a range in dielectric constant of more than about two to one. The corresponding range of wire-to-pin capacitance is even less; so insulator material may be said to have a relatively small influence on E.

4. *Influence of Pin Spacing.* Item E is naturally expected to be influenced by the pin spacing. However, the data bearing on this effect are too meager to report.

Item F—Losses Due to Unbalanced Displacement Currents Flowing in External Impedances such as Crossarms, Poles, Etc.

1. *General Characteristics.* In general characteristics, F is very similar to E. As E is caused by a displacement current which flows directly from one line wire to the other via the crossarm in the manner already discussed, so F is similarly caused by unbalance displacement currents flowing through crossarms, poles, etc. There is not sufficient space in this paper either to present details of the theory of these losses or to describe the many interesting tests made to illustrate the effect.

In brief, F is due, first, to any difference that may exist between the capacitance of the insulator on one wire and that on the other wire, and second, to other unbalances in capacitance such, for example, as those caused by the presence of other wires.

The first of these causes will be recognized to be very similar in nature to a second order effect of E and is accordingly small in magnitude, at least if the same kind of insulator is employed on each wire; so F, due to this particular cause, is normally small.

The second source is the more important one. F, resulting from it, is greatest in dry weather, like E. Here, F's importance is so great that transpositions in the insulator test line were found necessary, despite the line being only 200 feet in length. The dry-weather leakage of many of the insulators under test is so small that, without transpositions, errors caused by the presence of other wires have amounted to several hundred per cent.

The variation of F with frequency is much the same as E.

Also, like E, F is less in wet than in dry weather.

The general remarks made in the discussion of Item E regarding the influence of insulator design and material apply also to F.

For the well transposed lines used for carrier circuits and for the reasonably well balanced insulator capacitances that the standard construction gives, F contributes little to the total wet-weather leakage.

Item G—Displacement Currents Flowing Over Insulator Surfaces through High Resistance

1. General Characteristics. Over the carrier range of frequencies this item is the most important source of leakage in wet weather. It probably contributes more than all other sources combined. On this account, it may not be amiss to repeat here an already known theory which fits the results of the present study fairly well.

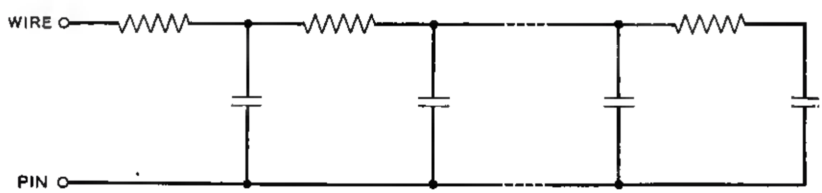


Fig. 21—Electrical equivalent of an insulator for the production of (G).

Consider an insulator, the outside of which is wet. Divide this surface into elements of area and take, for example, one of them near the bottom of the insulator. Assume a small displacement current to flow from the pin to this element through the small capacitance which exists between them. Now for this current to reach the wire, it must flow through the thin film of moisture lying between the element chosen and the wire. This film offers a high resistance to the current, not high enough, however, to seriously limit the current, but, nevertheless, sufficient to produce heat losses. These losses when integrated over the entire insulator surface become important and qualitatively, at least, account for the characteristics of item G.

Fig. 21 shows in much simplified form an electrical equivalent of this action. An inspection of this figure will throw some light on the subject.

For one thing, it is clear that the apparent wire-to-pin capacitance will decrease with increasing frequency. Wet-weather tests invariably show this effect. (Incidentally, this effect tends to reduce the magnitude of C and D, at the higher frequencies.)

It is also apparent that the conductance of this simplified circuit is zero at zero frequency and a maximum at very high frequencies. In the intervening range the conductance at first rises nearly as the square of the frequency, then the relation becomes more nearly linear and finally tapers off to a final constant maximum value.

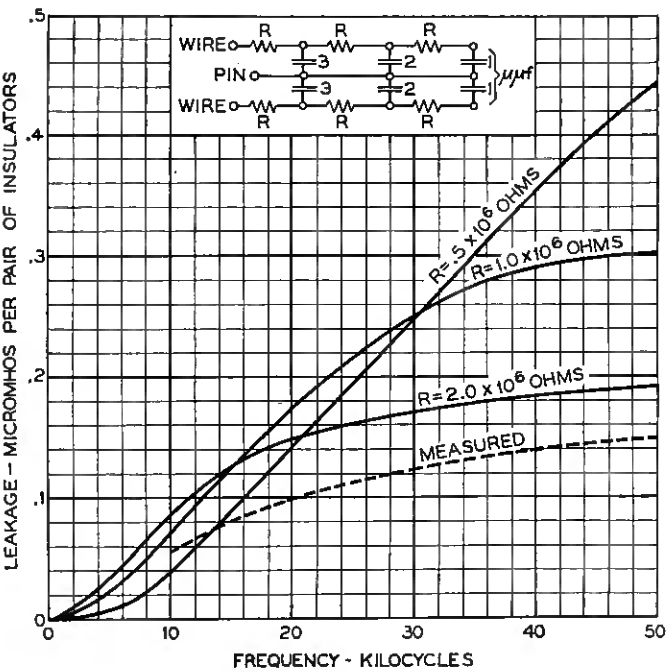


Fig. 22—Calculated and measured value of (G) for C. P. insulator.

If the distributed capacitance of an insulator were accurately known, as well as the resistivity of the surface film, the magnitude of G could be calculated. Unfortunately, both factors are difficult to determine with precision.

However, it is useful to carry out such a calculation, even though the factors be not accurately known.

Fig. 22 shows the variation of G with frequency, as calculated for the assumptions indicated on the drawing. The assumed distributed capacitance and surface resistivity are intended to represent those of the insulator of Fig. 6. A measured value of G for this insulator is shown by the dotted curve. The agreement is not very good, being no more than qualitative. However, the accuracy of the assumptions is not sufficiently high to expect a much better agreement.

In accordance with this theory, it is evident that G could be produced on a dry insulator by coating the surface with a thin metallic film of high resistance. This has been checked experimentally. The dry-weather conductance of an insulator of given shape and material can readily be increased ten or more times by such a coating. In its general magnitude this increase corresponds to that of the uncoated insulator between dry and wet weather.

Referring again to Fig. 21, G will be seen to be zero when the resistivity of the circuit is zero. Thus, if a metal coating of low resistance be placed on the outside surface, the value of G would be made negligible for that surface. However, such a coating causes the losses on the inner surfaces to increase. The net result is usually a substantial reduction in leakage, as shown by many tests comparing coated with uncoated insulators.

2. Influence of Insulator Design. From the foregoing discussion, the importance of insulator capacitance is quite apparent. For G to be small, it is not only desirable that the total capacitance shall be small, but particularly that any capacitance remote from the wire groove be small, unless the resistance of the path from wire to that capacitance can be maintained at a very high value.

Insulator design is important from this standpoint. It is also important because of its relation to the resistivity of the surface film, as was discussed in some detail under B.

3. Influence of Insulator Material. In general, the influence of insulator material has been less than that of design, especially after the insulator has aged for a couple of years. The influence of material on G is determined both by the dielectric constant of the material and its surface resistivity.

While the dielectric constant of the materials tested varies about two to one, the corresponding range of G is very much less, on account of the low dielectric constant of the air which occupies a part of the dielectric path.

The factor of surface resistivity is probably the more important one. For example, its effect on G probably accounts for most of the increase in leakage at high frequencies that accompanies aging, which is a very substantial one.

As to the relative importance of these two factors, tests have been made on only one design. At this writing the tests register a slight balance in the favor of a good surface over a good (low) dielectric constant.

The influence of material on G, while not great, gives the chief justification from the electrical standpoint for employing the better but more costly glasses.

Conclusion

Description of New Insulators. Two new insulators are available for use on carrier circuits of the Bell System. These replace the standard D. P. insulator of Fig. 1 in those cases where the additional expense of the new insulators is justified.

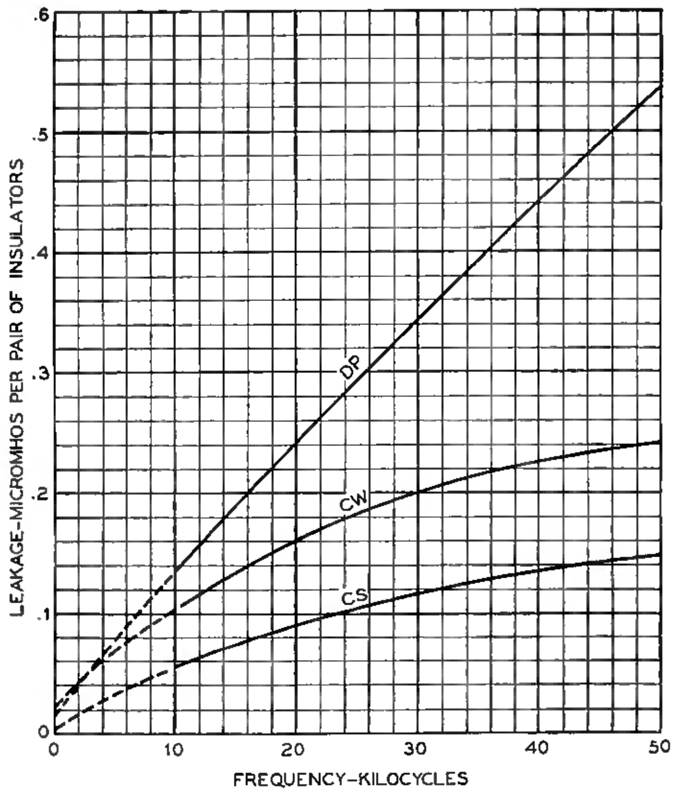


Fig. 23—Relative leakage of D. P., C. W. and C. S. insulators as measured at Phoenixville, Pennsylvania, in moderate rain.

One of the new types is designed for use on the wood pins of existing lines. It is known as the C. W. insulator and its design is shown in Fig. 3. The pair of pin thimbles illustrated at the bottom of that figure is first placed over the two wood pins, then the insulators are screwed into place over the thimbles, forcing the latter well into the threads of the wood. The thimbles are constructed of thin copper and are bonded together by a tinned copper strip.

The other new type is designed to screw directly over a steel pin. This is known as the C. S. insulator and its design is shown in Fig. 12. At each crossarm the two steel pins are bonded by means of a wire underneath the arm.

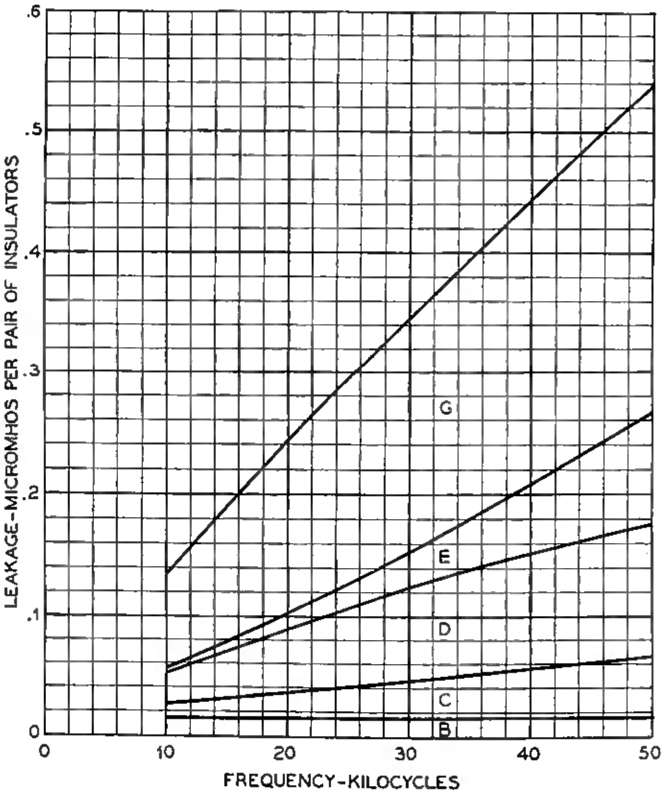


Fig. 24—Estimated allocation of leakage for D. P. insulator.

Both new designs are molded from borosilicate glass. This glass is more expensive than the alkali glass used in the old D. P. design. On this account, and on account of the pin thimbles in one case and the steel pins in the other, the new insulators cost more to install than did the old ones.

Both of the new designs were brought out several years ago before this study had conclusively demonstrated the importance of surface losses (item G). It will, therefore, be of some interest to discuss their performance in the light of the more complete knowledge.

Performance of D. P., C. W., and C. S. Insulators. The leakage of these three types, as measured on the insulator test line at Phoenixville, Pa., in a moderate rain, is given in Fig. 23. This measurement does not give a true picture of the relative efficiency of the three types because no two of them are aged alike. Besides, the relative efficiency varies considerably with different weather conditions. However, the measurement will serve our present purpose, which is to analyze the total leakage of each design and thus give a perspective which could not well be brought out in the detailed discussion of the several sources of

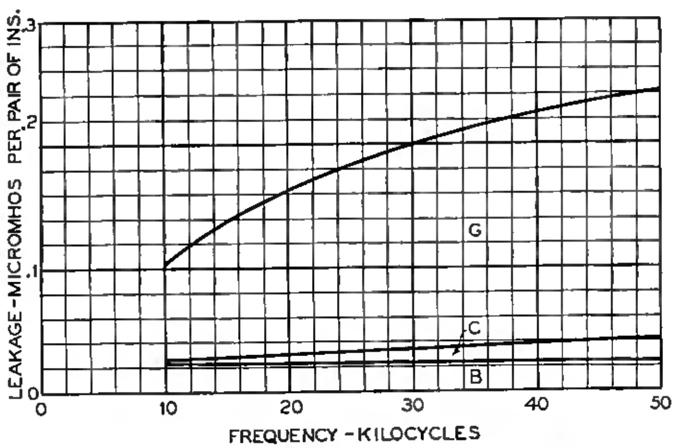


Fig. 25—Estimated allocation of leakage for C. W. insulator.

leakage. It should be pointed out and emphasized that the allocation of the total leakage to its component parts can be only very approximate.

Fig. 24 shows an estimate of the leakage contributed by the several sources for the D. P. design. The leakage directly through the insulator material is negligible and item A, therefore, does not appear. Similarly, the leakage due to unbalanced displacement currents flowing in crossarms, poles, etc., is considered negligible and, therefore, item F does not appear.

At a frequency of 30 kc., for example, B, the direct surface leakage or d.-c. leakage, is about 5 per cent of the total. The dielectric absorption in the glass C is about 10 per cent. The dielectric absorption in the wood pins D is about 20 per cent. The crossarm losses E contribute about 10 per cent and, finally, the losses on the insulator surfaces G contribute about 55 per cent.

Fig. 25 shows a similar estimate for the C. W. design. Here the bonded pin thimbles shield the wood pins from any electric field and thus eliminate dielectric absorption D from the pins. Similarly, by short-circuiting the crossarms, the losses occurring there are eliminated. Accordingly, both items D and E are made negligible and do not appear.

Of the remaining factors, the direct surface leakage B contributes about 12 per cent of the total (at 30 kc., for example). The losses in the glass C are liberally estimated at about 5 per cent. Finally, the surface losses G contribute over 80 per cent of the total.

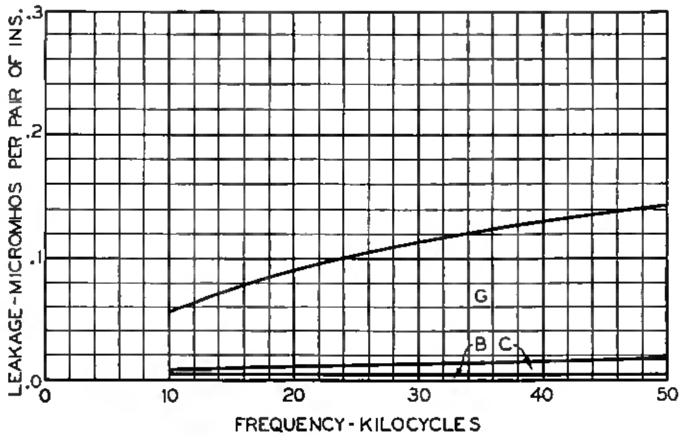


Fig. 26—Estimated allocation of leakage for C. S. insulator.

Comparing the C. W. performance in this test with that of the D. P., we find that most of the improvement shown by the C. W. has resulted from the elimination of items D and E. Due to the single skirt design of the C. W. and the pin thimbles, item B has been increased in magnitude. The loss in the glass C has been decreased by the use of borosilicate glass. However, the most important item G has been only slightly reduced, and if C. W. insulators in this test were aged as far as the D. P. the C. W. might show no improvement with respect to G. The pin thimble construction tends to increase the insulator capacitance and thus tends to make G larger for the C. W. than for the D. P. design. The use of a borosilicate glass with its lower dielectric constant counteracts this action somewhat.

The estimated division of losses for the C. S. design is given in Fig. 26. The use of metal has eliminated any dielectric absorption D from occurring in the pins. The bonding of the pins by wire has eliminated crossarm losses E.

Of the remaining factors, the direct surface leakage *B* contributes about 4 per cent or less of the total at 30 kc. The losses in the glass *C* are liberally estimated at 10 per cent or less, while the surface losses *G* contribute about 85 per cent or more.

In this design the absolute magnitude of *B* has been decreased somewhat, chiefly because the small diameter of the steel pin permits a small diameter of insulator, the advantages of which were pointed out in the detailed discussion of this item.

The low capacitance made possible by the small steel pin has helped to make both *C* and *G* relatively small, although most of the improvement in *C* is due to the borosilicate glass.

The improvement in the surface losses *G* over the D. P. design is quite marked.

For the new designs the two factors *B* and *G* are the controlling ones. In this particular test *B* happens to be quite small in magnitude, and would naturally lead one to conclude that *B* had been made unnecessarily small at the expense of *G*, especially since these two items in many respects place conflicting requirements on insulator design. However, B has been observed at times to reach a value as high as one third of the total leakage at 50 kc. The necessity of engineering for such cases makes the design more reasonable, especially when it is recalled that the insulators must serve for direct current and low frequencies, as well as for the carrier range.

Of the new designs, the electrical superiority of the C. S. over the C. W. design is apparent. This fact, together with economic considerations, has led to the almost universal choice of the C. S. rather than the C. W. type for the field of application of the new insulators in the telephone plant.

The utility of the C. S. insulators in the telephone plant will be more clearly apparent from a consideration of the reduction in attenuation which their use brings about.

The losses in transmission over a pair of wires at carrier frequencies come chiefly from two sources: one, substantially fixed in magnitude, depending mainly on the resistance of the wires; the other, quite variable in magnitude, depending on the leakage conductance between the wires and, therefore, on the weather.[6]

In the case of 165-mil copper wires on 12-inch spacing these two components of loss are approximately equal in wet weather at 30 kilocycles when the older type of insulators are used. The C. S. type cuts

[6] For a more detailed discussion of attenuation see a companion paper, "The Transmission Characteristics of Open-Wire Telephone Lines," by E. I. Green, presented at the Summer Convention of the A. I. E. E., Toronto, June, 1930 and printed in this issue of the *Bell System Technical Journal.*

this variable component about in half at that frequency thus reducing the total wet-weather attenuation of these wires to about 75 per cent of its former value. For smaller sizes of wires the percentage reduction is correspondingly less.

The benefits of the lesser attenuation can be utilized in the plant in various ways, depending on local conditions; for example, in increasing repeater spacing, in employing smaller gages of wires, or in increasing the number of insulators per mile to provide for better transposition designs.

In addition, the new insulators, in having reduced the variable component of loss, improve the stability of carrier circuits to a marked degree.

Acknowledgment

Only the electrical features of the new designs have been discussed. Closely related to these are the many mechanical problems which naturally arise in new construction. These latter problems, during the development of the new designs, came under the supervision of Mr. C. S. Gordon, assisted by Mr. J. T. Lowe.

The Corning Glass Works has cooperated in molding special experimental insulators of various compositions.

Data on the electrical properties of numerous glass compositions have been supplied by the Bell Telephone Laboratories.

The writer desires to express his thanks to Messrs. F. A. Leibe, L. R. Montfort and L. Staehler for assistance in the measurements and to Mr. H. R. Nein for assistance in the preparation of this paper.

The Transmission Characteristics of Open-Wire Telephone Lines[1]

By E. I. GREEN

Values of the primary transmission constants R, L, G, and C for open-wire telephone lines are presented, and the factors which affect these constants in practise are discussed. Consideration is then given to the constants which are of principal interest in telephone work, namely, the attenuation, the characteristic impedance, the phase constant, and the velocity of propagation. Data regarding these characteristics are given for the frequency range from 0 to 50,000 cycles.

NEARLY 3,000,000 miles of open wire are now furnishing toll service in the Bell System, and this total is increasing at a rate of more than 100,000 miles a year. Hence, the subject of the transmission characteristics of open-wire circuits, in addition to being of considerable natural interest, is of no little importance in many branches of telephone work. In the design of apparatus to be associated with the open-wire circuits as well as in the engineering and maintenance of the facilities derived from them, a knowledge of these transmission characteristics is indispensable.

The problem of determining the characteristics of the open-wire circuits has, of course, been coexistent with the circuits themselves, and hence dates back to the beginnings of telephony. Of late years, however, there has been a very decided change in the nature and scope of the problem. This has resulted from many factors, particularly (*a*) the extensive application of carrier telephone and telegraph systems[2] and (*b*) the constantly increasing length of the long distance circuits. The first of these factors has extended the transmission range upward from about 3000 cycles to about 30,000 cycles, and may well extend it higher in the future. The second, in combination with the higher standards which are now applied in long distance transmission, has required greater accuracy in the data, emphasizing especially the importance of time and space variations in the characteristics. Also, recent changes in the construction of open-wire lines (to be described later) have necessitated substantial additions to the data.

[1] Presented at the Summer Convention of the A. I. E. E., Toronto, Ont., Canada, June 23–27, 1930.

[2] See "Carrier Systems on Long Distance Telephone Lines," by H. A. Affel, C. S. Demarest, and C. W. Green, *A. I. E. E. Trans.*, Vol. 47, 1928, pp. 1360–1386. *Bell System Tech. Jl.*, July 1928, pp. 564–629.

There will be studied in this paper those inherent characteristics of open-wire lines which are used most frequently in telephone transmission work. These characteristics are: first, the attenuation, second, the impedance, and third, the phase characteristic, with which must be coupled its near relative, the velocity of propagation. The range of frequencies to be covered is fixed on the one hand by the d.-c. telegraph systems, as well as by the program transmission circuits, whose lower frequencies extend to 100 cycles or less, and on the other hand by the carrier telephone and telegraph systems, which make the range up to about 50,000 cycles of interest.

Line Construction Arrangements

In order to study the characteristics of open-wire lines it is necessary to know something of the constructional arrangements which are employed. The conductors most commonly used for the open-wire tele-

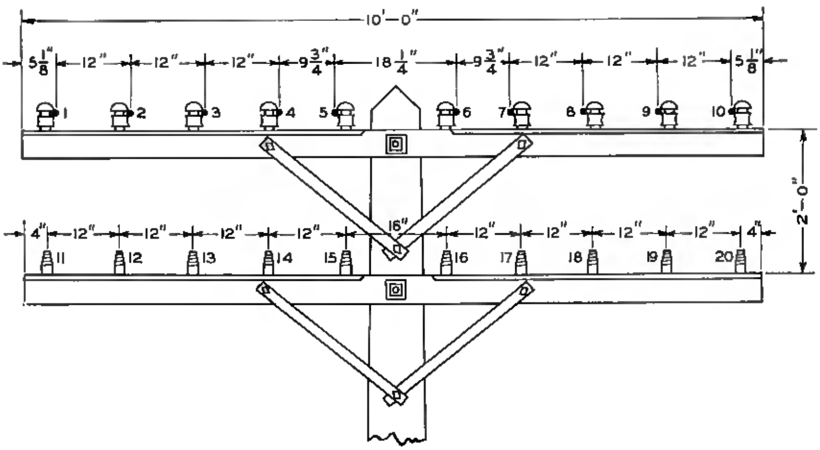


Fig. 1—Configuration of an open-wire line with 12-in. non-pole pairs.

phone lines of the Bell System are of 165-mil (No. 8 B. W. G.), 128-mil (No. 10 N. B. S. G.), and 104-mil (No. 12 N. B. S. G.) hard-drawn copper. These are the conductors usually employed for carrier systems. Other gages of copper, as well as a small amount of iron or steel wire, are used to some extent for voice-frequency and d.-c. telegraph transmission only.

The wires of the lead (as an open-wire line is frequently designated in telephone parlance) are strung on poles, the normal spacing and numbering of wires being generally as shown in Fig. 1. Starting at the left end of the crossarm, the adjacent horizontal wires are grouped in pairs, wires 1 and 2 comprising one pair, 3 and 4 another, etc. The characteristics of the pairs are of primary interest. Phantom circuits, which are derived from two pairs or side circuits, will be discussed later. The two wires of each pole pair (that is, a pair which bestrides the pole) are about 18 in. apart, and those of each non-pole pair 12 in. apart.

There has recently come into vogue a different arrangement of wires which is designed to reduce the coupling between circuits and thus permit a maximum use of carrier facilities. This arrangement is portrayed in Fig. 2. In this newer configuration the separation between the wires of each non-pole pair is reduced to 8 in., and the horizontal separation between pairs is widened to 16 in.

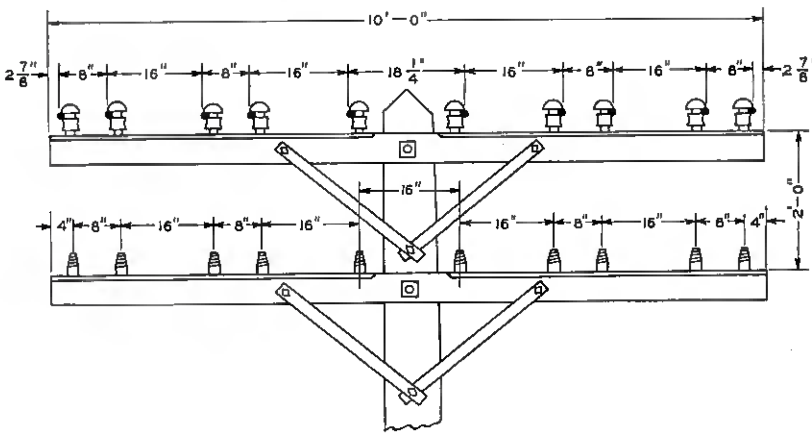


Fig. 2—Configuration of an open-wire line with 8-in. non-pole pairs.

The ordinary spacing between poles on open-wire toll lines is 132 ft., corresponding to a total of 40 poles per mile. Where additional strength is required, the number of poles per mile may be as high as 50, while outside of the heavy sleet area it may be as low as 30.

The types of insulators employed on open-wire lines will be discussed under the heading of leakage conductance.

Two methods of transposing the wires are in current use. In the older of these, which is illustrated in Fig. 3, the wires are brought at the transposition pole to a "drop bracket." The transposition is accomplished over a total distance of two spans by gradually rotating the plane of the wires through 180 deg. It will be seen that the wire configuration is abnormal throughout the two spans. In the newer method, the wires are crossed practically at a point. This may be done by means of two brackets known as "break irons," as illustrated in Fig. 4, or by means of a single bracket. The "point" transpositions preserve the nominal spacing between wires and thus avoid the irregularities in spacing which occur when drop bracket or "rolling"

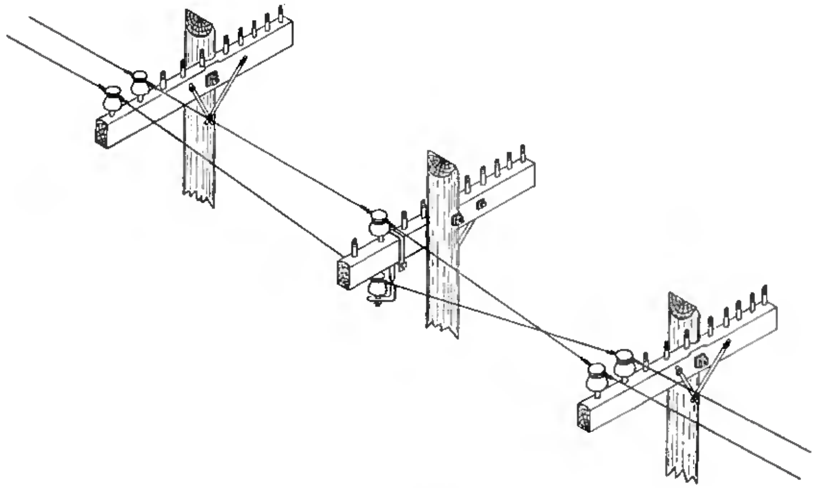


Fig. 3—Transposition of wires with drop bracket.

transpositions are used. With the point transpositions, however, two pairs of insulators are required at each transposition point as com-

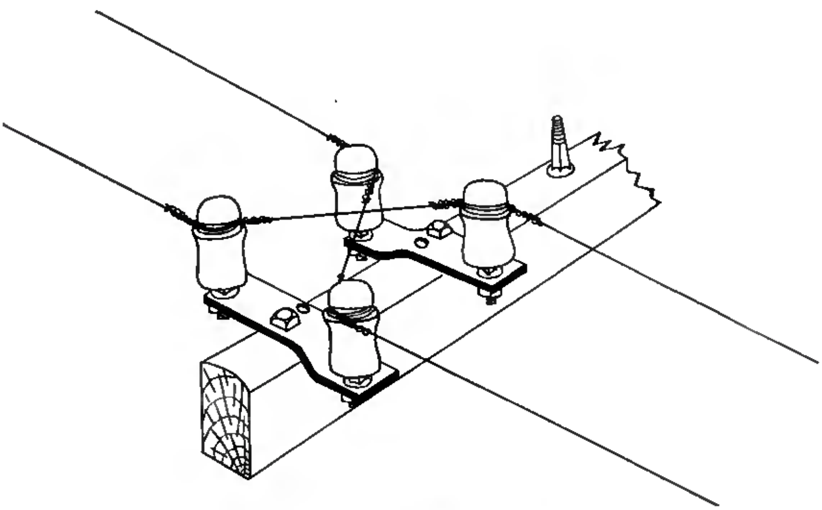


Fig. 4—"Point" transposition on break irons.

pared with a single pair of insulators at a non-transposition point. This results in an increase in the total number of insulators and in variations in the number of insulators on different pairs because of the different number of transpositions employed.

Primary Constants

It should be noted here and now that the phenomena of line transmission are the same throughout all of the frequency range under consideration. Transmission over wires at high frequencies is accomplished in precisely the same manner as transmission at low frequencies, the wires acting as the guiding medium for the energy in both cases, and the same theory may be applied to both.

A review of the well-known theory for the propagation of alternating currents over wires will show that the line characteristics in which we are interested are dependent upon the four quantities known as the primary constants of the circuit. These are as follows:

R = Series resistance in ohms per mile.
L = Series inductance in henries per mile.
C = Shunt capacitance in farads per mile.
G = Shunt leakage conductance in mhos per mile.

These quantities may be stated per mile of wire or per mile of circuit. In this paper all values will be per mile of circuit, or, as it is commonly expressed, per loop mile.

Unfortunately the constants R, L, G, and C are by no means constant in practise. Indeed there could scarcely be a more fickle set of quantities. They are subject to change by a great variety of factors, of which the most important is, of course, the frequency. Hence it is evident that in order to determine the practical values of the attenuation, impedance, and velocity for open-wire circuits, we shall have to examine the behavior of the primary constants, R, L, G, and C.

Resistance

First in the list of primary constants is generally named the conductor resistance. The method of computing the d.-c. resistance is well known and requires no explanation here. In such computations it is assumed that the current density is uniform throughout the cross-section of the conductor. With alternating current, however, the familiar phenomenon of skin effect tends to produce a non-uniform current distribution, and hence to increase the resistance. If the two wires of a circuit are close together, the effective a.-c. resistance of each wire is likewise increased by the presence of the parallel conductor, due to what is known as proximity effect. In cable conductors, especially when used for carrier frequencies, proximity effect is very important, but it is negligible in open-wire circuits because of the large separation between wires.

The method of determining the skin effect resistance of round wires is presented in various publications, and the theoretical results have been experimentally confirmed on numerous occasions.[3] Values of the

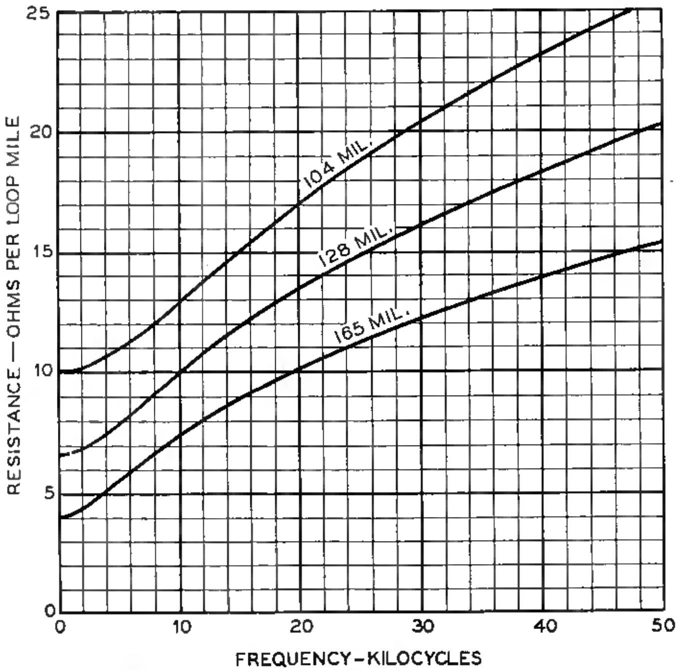


Fig. 5—A.-C. resistance of open-wire pairs at 20 deg. cent. (68 deg. fahr.).

a.-c. resistance of 165, 128, and 104-mil copper pairs at 20 deg. cent., (68 deg. fahr.) determined in accordance with skin effect theory, are plotted in Fig. 5. It will be noted that the increase in resistance due to skin effect is small in the voice range, but rather astoundingly large in the carrier range, amounting at the higher frequencies to from 200 to nearly 400 per cent.

Experimental evaluations of open-wire resistance are in extremely

[3] See "Wave Propagation Over Parallel Wires—The Proximity Effect," J. R. Carson, *Phil. Mag.*, Vol. 41, April 1921, pp. 607–633; "Experimental Researches on Skin Effect in Conductors," A. E. Kennelly, F. A. Laws, and P. H. Pierce, *A. I. E. E. Trans.*, Vol. 34, Part 2, 1915, pp. 1953–2021, and "Skin Effect Resistance Measurements of Conductors at Radio Frequencies," A. E. Kennelly and H. A. Affel, *I. R. E. Proc.*, Vol. 4, No. 6, Dec. 1916, pp. 523–574.

close agreement with the values given in Fig. 5.[4] For old wires, however, the resistance may be somewhat higher than these values. This is because of the presence of contact resistance in the twisted sleeve joints in the wires and also perhaps because of an actual decrease in the conductor diameter occasioned by corrosion. The increase in the d.-c. resistance of old wires due to these causes may be as much as 5 per cent. The corresponding percentage of increase in the a.-c. resistance will, of course, be much smaller. The d.-c. resistance of a copper wire varies with temperature according to the familiar formula:

$$R_0 = R_{01} [1 + \alpha_1 (t - t_1)], \qquad (1)$$

where R_0 and R_{01} represent the d.-c. resistance at temperatures t deg. cent. and t_1 deg. cent. respectively, and α_1 is the d.-c. temperature coefficient of resistance of copper at t_1 deg. cent. At a temperature of 20 deg. cent. the value of α_1 is generally taken as 0.00393.

Similarly, the a.-c. resistance R of a copper wire at a temperature t may be represented as follows:

$$R = R_1 [1 + A_1 (t - t_1)], \qquad (2)$$

where R_1 = a.-c. resistance at temperature t_1 deg. cent.,

$A_1 = \dfrac{1}{R_1}\dfrac{dR}{dt}$ = a.-c. temperature coefficient of resistance of copper at t_1 deg. cent.

Now the skin effect resistance ratio depends upon the magnitude of the d.-c. resistance, being smaller the larger the resistance. Hence, a given change in temperature which changes the d.-c. resistance produces a change in the opposite direction in the skin effect resistance ratio, so that the percentage change in the a.-c. resistance is less than the percentage change in the d.-c. resistance. In other words, A_1 is less than α_1. As illustrated in Fig. 6, the a.-c. temperature coefficient of resistance for open-wire pairs, starting at the d.-c. value α_1, straightway decreases as the frequency is increased, and at high frequencies approaches a value of $\dfrac{\alpha_1}{2}$. An explanation for this asymptotic value is presented in Appendix I.

The temperature assumed by the conductors of open-wire lines depends of course, upon the weather conditions which prevail in different sections of the country. In order to obtain information on this subject,

[4] The value of R, and also that of the other primary constants, may be determined directly from open and short-circuit impedance measurements on a line short enough to avoid propagation effects. A longer line may be used instead, in which case it is necessary to correct for such effects.

a study has been made of the Weather Bureau records for a number of representative cities in various parts of the country. The chief interest naturally centers in the extreme temperatures reached by the wires. It appears that on the average the air temperature will not drop below about − 20 deg. cent. (− 4 deg. fahr.) on more than about 10 days per year in the colder sections of the country, while a limiting temperature of about 35 deg. cent. (95 deg. fahr.) will not be exceeded on more than about 10 days per year in the warmer sections of the country. Because of imperfect radiation, the temperature of a wire in

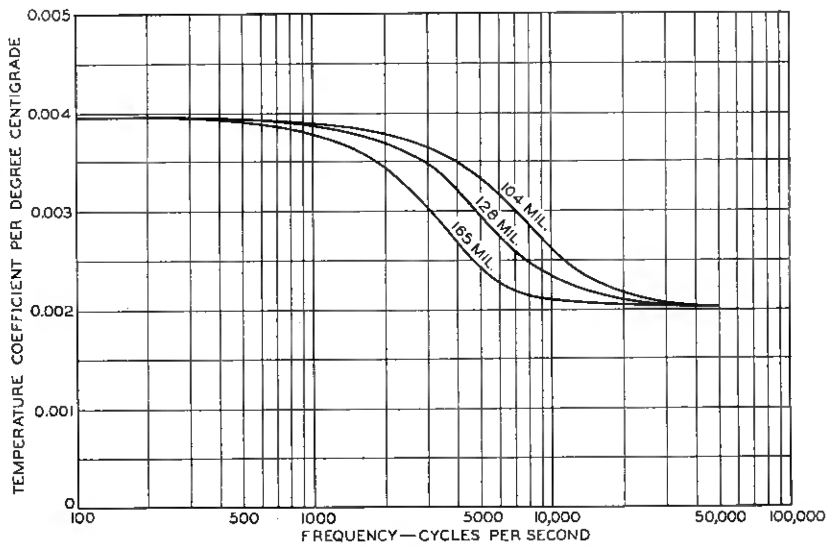


Fig. 6—A.-C. temperature coefficient or resistance for open-wire pairs at 20 deg. cent.

the sun will ordinarily exceed the temperature of the surrounding air by a small amount. A few tests have indicated that for open wires the increase over the air temperature on a warm day is not more than 5 deg. cent. Temperatures of − 20 deg. cent. (− 4 deg. fahr.) and 40 deg. cent. (104 deg. fahr.), therefore, appear to be representative values for the limiting temperatures assumed by open-wire lines. Reference to Equation (1) shows that this range of temperature gives possible variations in the d.-c. resistance of 16 per cent below and 8 per cent above the value for 20 deg. cent.

The total annual change in resistance at any one place will, of course, be less than the sum of the above changes. In the Middle West, however, where the weather variations are much greater than in other parts of the country, the total annual change in d.-c. resistance may be as much as 20 per cent. This section of the country has also the greatest diurnal range of temperature, giving a d.-c. resistance variation of as much as 8 per cent.

Inductance

The inductance of a circuit formed of two parallel wires whose distance between centers is negligible compared with their length is

$$L = 0.64374 \left[2.3026 \log_{10} \frac{2D}{d} + \mu\delta \right] \times 10^{-3} \text{ henrys per loop mile,} \quad (3)$$

where the diameter of each wire d, and the distance between their centers D, are expressed in the same units, where μ is the permeability, and δ is a factor depending on the frequency.

The tendency of alternating currents to concentrate on the surface of a wire evidently reduces the magnetic flux within the wire and decreases the internal inductance of the wire. In Equation (3) the internal inductance is represented by the factor $\mu\delta$. At low frequencies, for which the current is uniformly distributed across the cross-section of the wire, the value of δ is 0.25. For very high frequencies there is practically no magnetic flux within the wire, and the value of δ is zero. Between these frequency limits the value of δ is determined with the aid of skin effect formulas or tables.[5] For the wire diameters and spacings employed on open-wire lines the change in the total inductance due to skin effect is relatively small.

It is assumed in Equation (3) that the two wires are suspended in space or at a considerable distance from the ground and from other wires. In practise, the presence of other wires probably has some effect on the inductance, but for well transposed lines this effect is negligibly small.

The inductance at different frequencies of 165, 128, and 104-mil copper pairs having various spacings between wires is shown in the following table.

As will be seen from Equation (3), the inductance varies with the logarithm of the separation between wires. In the table the values of inductance are shown for pole pairs, which have a separation between wire centers of about 18.25 inches, and for non-pole pairs having wire separations of 12 and 8 inches. The values of inductance given in the table have been closely checked by measurements on open-wire pairs.

[5] See, for example, Circular No. 74 of the Bureau of Standards, "Radio Instruments and Measurements."

Frequency-cycles per second	Inductance of open-wire pairs—henrys per loop mile								
	165-mil			128-mil			104-mil		
	8-in.	12-in.	18.25-in.	8-in.	12-in.	18.25-in.	8-in.	12-in.	18.25-in.
0	0.00311	0.00337	0.00364	0.00327	0.00353	0.00380	0.00340	0.00366	0.00393
1,000	0.00311	0.00337	0.00364	0.00327	0.00353	0.00380	0.00340	0.00366	0.00393
10,000	0.00305	0.00331	0.00358	0.00323	0.00349	0.00376	0.00338	0.00364	0.00391
25,000	0.00301	0.00327	0.00354	0.00319	0.00345	0.00372	0.00334	0.00360	0.00387
50,000	0.00299	0.00325	0.00352	0.00317	0.00343	0.00370	0.00331	0.00357	0.00384
Infinite	0.00295	0.00321	0.00348	0.00311	0.00337	0.00364	0.00324	0.00350	0.00377

Capacitance

The capacitance of two parallel wires in space with a distance between centers which is negligible compared with their length is

$$C = \frac{0.019415}{\log_{10} \frac{2D}{d}} \times 10^{-6} \text{ farads per loop mile.} \tag{4}$$

It will be noted that the capacitance varies in inverse relation to the separation between wires.

As in Equation (3), it is assumed in this formula that the two wires are suspended in space or at a considerable distance from the ground and from other wires. On an actual line the capacitance of a pair is changed to an appreciable extent by the presence of other wires, and to a slight extent by the capacitance to ground. The true capacitance between the two wires under actual conditions may be derived from the direct capacitances between all wires and the direct capacitances of all wires to ground.[6] The capacitance is not changed to any great extent by skin effect.

The means of insulation and support provided at each pole have an appreciable effect on the capacitance of a pair of wires, especially in wet weather. This is due to the fact that the insulators and, under certain conditions, the pins and parts of the crossarms, act as the dielectric of small condensers which are, in effect, shunted between the line wires. These effects are being discussed in a companion paper.[7] The percentage increase in capacitance due to the insulators varies with different weather conditions and different types of insulators, ranging

[6] See Technical Report No. 54 of the Railroad Commission of the State of California, Joint Committee on Inductive Interference, entitled "Inductive Interference Between Electric Power and Communication Circuits," 1919.

[7] See "A Study of Telephone Line Insulators," by L. T. Wilson, printed in this issue of the *Bell System Technical Journal*.

from about 0.5 to 4 per cent of the capacitance between line wires.

Values of the capacitance of 165-, 128-, and 104-mil pairs in space and on a 40-wire line are given in the following table:

Wire spacing	Capacitance—microfarads per mile					
	165-mil		128-mil		104-mil	
	In space	On 40-wire line	In space	On 40-wire line	In space	On 40-wire line
8 in.	0.00977	0.00996	0.00926	0.00944	0.00888	0.00905
12 in.	0.00898	0.00915	0.00855	0.00871	0.00822	0.00837
18.25 in.	0.00828	0.00863	0.00791	0.00825	0.00763	0.00797

As before, values are given for pairs having wire separations of 8, 12, and 18.25 in. These values include an allowance for the dry weather capacitance of the insulators. The difference between the values in space and on a 40-wire line indicates the importance of the effect of the other wires, the insulators, etc., upon the capacitance. The capacitance values given in the table are fairly representative of the values that will obtain on well transposed lines.

The values of inductance and capacitance which have been given are based on the assumption that the nominal separation between wires is preserved throughout the entire line. This is not the case when drop bracket transpositions are employed. As has been pointed out, the wires are brought closer together at the drop brackets, thereby increasing the capacitance and decreasing the inductance. The amount of the change in inductance and capacitance due to this cause ranges from 1 to 5 per cent for the transposition arrangements designed for carrier system operation.

Leakage Conductance

The leakage conductance per unit length of circuit, which is represented in the transmission formulas by the symbol G, is by far the most erratic of the primary constants. Since it is a momentous factor in the attenuation its investigation has been very actively prosecuted over a considerable period of time.

The determination of the value of G for direct current is quite simple, involving merely a measurement of the actual conductance between wires for a length of circuit short enough to avoid propagation effects. For alternating currents, however, it is customary to employ an equivalent value of G which includes all of the losses suffered by the power transmitted over the pair except the normal I^2R loss in the wires themselves. This inclusion of numerous little-understood losses in the general term leakage has at times served to insulate the individual losses from analysis. Methods of determining the value of the "equivalent leakage conductance" and of analyzing its component losses are available, however.[8]

The nature and magnitude of the different losses which occur at the insulators are being discussed in detail in a parallel paper.[7] Accordingly, only a brief mention will be made in this paper of the types of insulators which are now in use on the open-wire lines of the Bell System, and of the values of leakage conductance experienced with these different types.

The *DP* or double-petticoat glass insulator illustrated in Fig. 7 is now standard for use on all important toll circuits, except those equipped with the special carrier insulators discussed below. On a number of older circuits single-petticoat glass insulators, known as toll insulators (see Fig. 7) and double-petticoat porcelain insulators are still in place.

In view of the numerous and complex sources of leakage loss, it is not surprizing that the leakage conductance for a given pair at a particular frequency varies with changing weather conditions and with the age of the insulators over a very wide range of values. Because of this wide range of variation it is possible to give here only selected leakage values which serve for engineering purposes. Values of the total leakage conductance for open-wire pairs equipped with *DP* insulators are plotted in Fig. 8. These values are intended to represent the highest values ordinarily obtained on an old circuit which is in a good condition of maintenance. The wet weather values have been so chosen that they should be exceeded on only a few days of the year, while the dry weather values represent the performance that should be expected from any circuit on a clear, dry day.

Particular difficulty is experienced in selecting standard values of d.-c. leakage owing to the fact that the range of values encountered in practise is exceedingly great. The measured values depend to a great extent on the degree to which the line is kept free from tree branches, foliage, moss, broken insulators, and other possible sources of leakage. These special sources of loss, of course, represent a much smaller part of the total leakage losses at carrier frequencies.

The standard values of leakage conductance are derived for a line having 40 pairs of insulators per mile. Where the number of insulators

[8] See "Methods of Measuring the Insulation of Telephone Lines at High Frequencies," by E. I. Green, *A. I. E. E. Trans.*, Vol. 46, 1927, pp. 514–519.

differs greatly from this figure, it is necessary to correct the leakage values accordingly. Differences in the number of insulators per mile result from the use of different types and numbers of transpositions, different pole spans, double crossarm construction, etc.

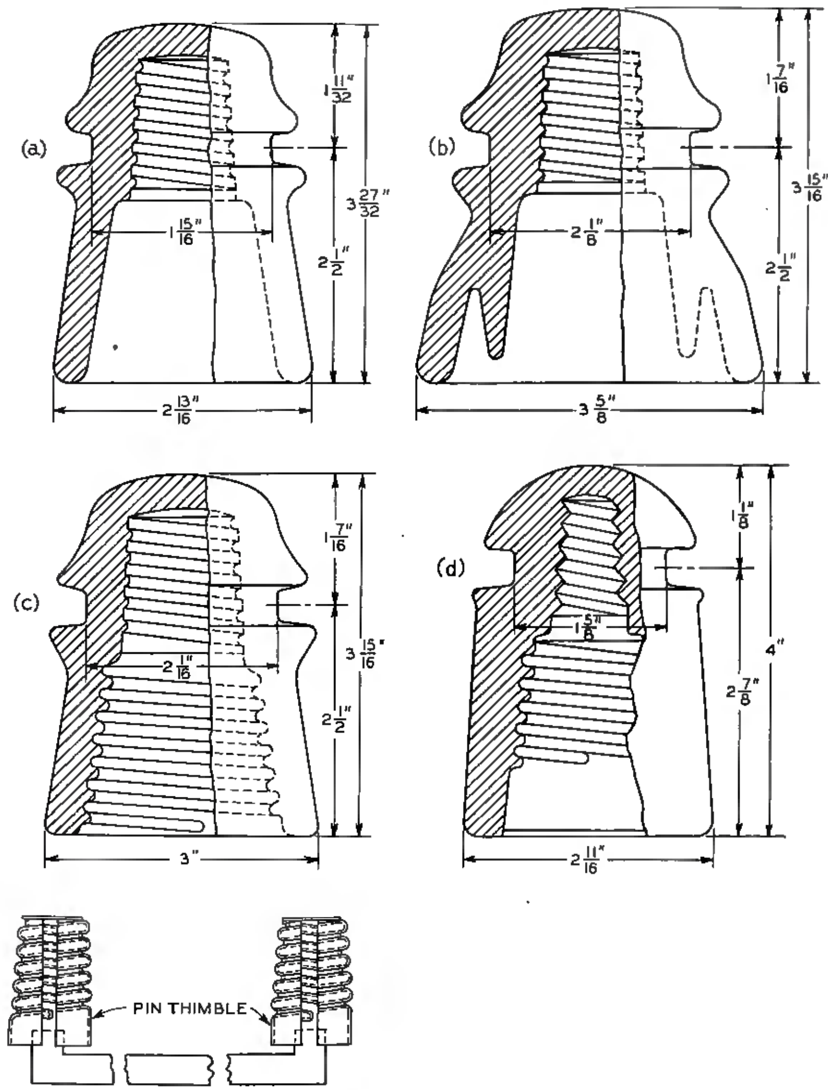


Fig. 7—Types of insulators employed in the Bell System.
a. Toll insulator.
b. *DP* insulator.
c. *CW* insulator and pin thimble.
d. *CS* insulator.

Considerable study has been given to methods of reducing the leakage conductance, particularly at carrier frequencies, and two new types of insulators have been developed for this purpose. In these there is used an improved dielectric (borosilicate glass) which has a low power factor and a reasonably low dielectric constant, as well as good chemical stability. Two expedients for eliminating losses in the pins and crossarm are employed. In the *CW* insulators, illustrated in Fig. 7, two

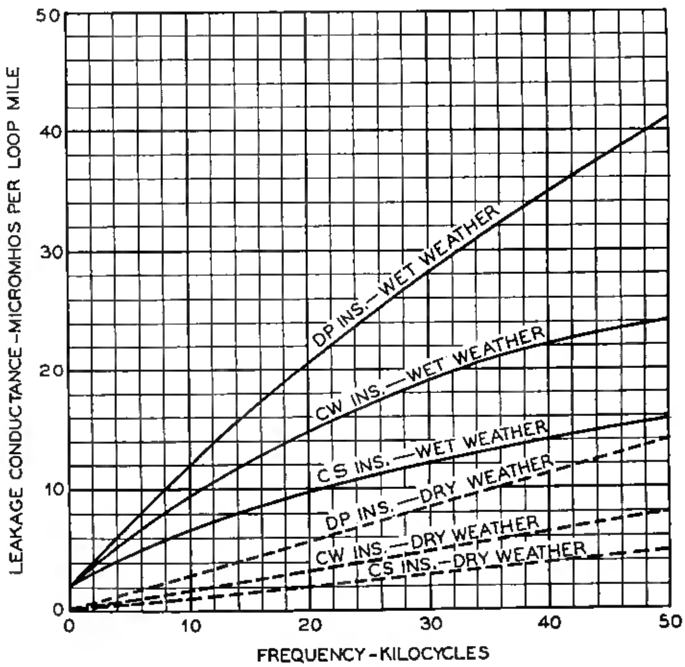


Fig. 8—Leakage conductance of open-wire pairs equipped with different types of insulators.

metal shells or thimbles are bonded together and placed over the wooden pins. In the *CS* insulator, also shown in Fig. 7, steel pins are employed and the two steel pins of the pair are bonded together. In the past few years *CS* insulators have been applied to the open-wire lines of the Bell System in increasingly large numbers.

With these two types of insulators a substantial reduction in the total leakage conductance is brought about. The best available figures for the limiting values of leakage conductance for these types are shown in Fig. 8. A further advantage obtained through the application of the new insulators is that of stabilizing the attenuation at carrier frequencies. Experience has indicated that the *CS* and *CW* insulators reduce the daily leakage (and attenuation) variations due to change in weather conditions to about one-third and one-half, respectively, of their value for *DP* insulators. This degree of stabilization is not indicated by the differences between the dry and wet weather leakage values shown in the figure, but it must be recalled that these values represent extreme conditions, while the stabilization referred to above is for average conditions.

Attenuation

The attenuation constant is the real part α of the propagation constant γ as given in the familiar formula

$$\gamma = \alpha + j\beta = \sqrt{(R + jL\omega)(G + jC\omega)}. \tag{5}$$

The attenuation constant is also given by the following expression

$$\alpha^2 = 1/2\ [\sqrt{(R^2 + L^2\omega^2)(G^2 + C^2\omega^2)} - (LC\omega^2 - RG)]. \tag{6}$$

Where $L^2\omega^2$ is large compared to R^2 and $C^2\omega^2$ is large compared to G^2, it can be shown[9] that Equation (6) reduces to

$$\alpha \doteq \frac{R}{2}\sqrt{\frac{C}{L}} + \frac{G}{2}\sqrt{\frac{L}{C}}. \tag{7}$$

This formula is very useful for computing the attenuation of open-wire circuits at carrier frequencies, in which case its accuracy is adequate for all practical purposes. It is frequently of value also for quick computations of the approximate attenuation of open-wire circuits in the voice range.

The first term of Equation (7) represents the series losses, and is commonly referred to as the "resistance component of attenuation," while the second term represents the shunt losses, and is called the "leakage component of attenuation." It will be observed that the resistance component of attenuation varies inversely with the quantity $\sqrt{\frac{L}{C}}$, while the leakage component varies directly with the same quantity. This quantity $\sqrt{\frac{L}{C}}$, as will be seen later, represents the nominal characteristic impedance of the circuit.

It is shown in Appendix II that a circuit of fixed resistance and leakage conductance will have minimum attenuation when the ratio of L to

[9] See "Transmission Circuits for Telephonic Communication," by K. S. Johnson, N. Y., Van Nostrand, 1927.

C is such as to make the resistance component equal to the leakage component. Because of the variation of the resistance and leakage conductance with frequency, the ratio of L to C which gives minimum attenuation evidently depends upon the frequency. At voice frequencies the resistance component for open-wire circuits is, as a rule, considerably larger than the leakage component, so that it is generally possible to reduce the voice-frequency attenuation by inserting loading coils, which increase the value of L and thus reduce the resistance component at the expense of an increase in the leakage component. The amount of reduction in attenuation obtainable by loading is

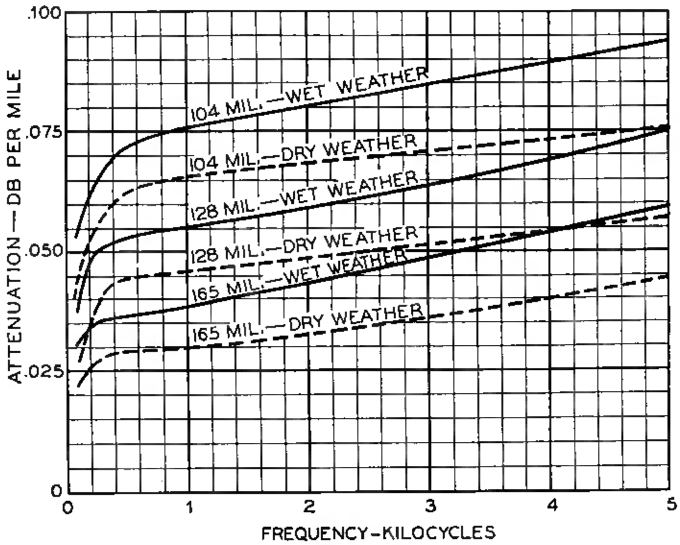


Fig. 9—Voice-frequency attenuation of open-wire pairs equipped with *DP* insulators.

evidently limited by the value of the leakage component and by the additional resistance which is contributed by the loading coils. For open-wire circuits at carrier frequencies the value of the leakage component of attenuation is quite large in comparison with the resistance component, and coil loading would, in general, be detrimental. At the present time the use of loading on the open-wire circuits of the Bell System has been practically abandoned. Owing to the importance of other factors, especially the line crosstalk, it is ordinarily impracticable to design the open-wire circuits to secure precisely the minimum attenuation at the highest working frequency. In the carrier frequency range, however, the wet weather attenuation of the pairs most commonly used is not materially higher than the theoretical minimum.

Values of the attenuation constant of open-wire pairs of different gages when equipped with *DP* insulators are presented in Figs. 9 and 10. The values are plotted in db per mile.[10]

The attenuation values of Figs. 9 and 10 have been determined for a temperature of 20 deg. cent. (68 deg. fahr.) and for the dry and wet

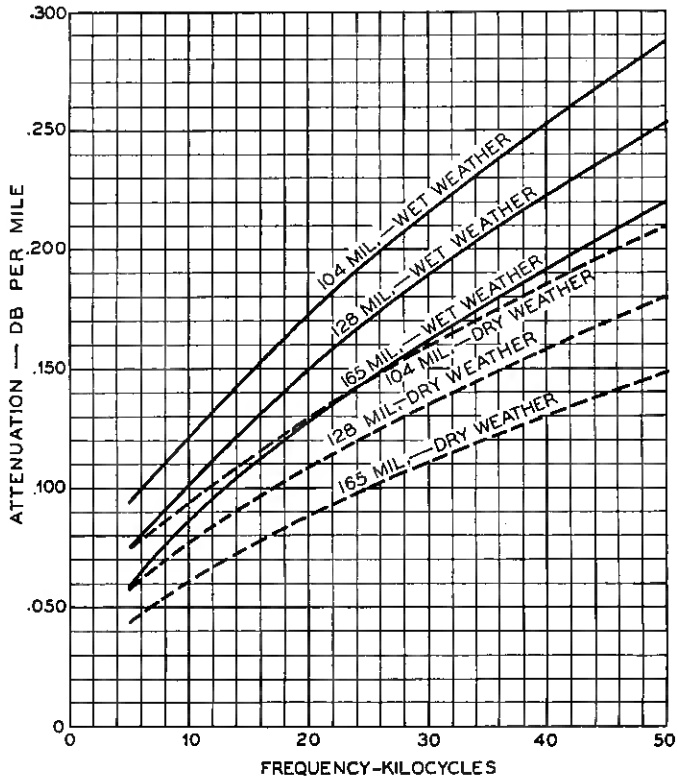


Fig. 10—Carrier frequency attenuation of open-wire pairs equipped with *DP* insulators.

weather values of leakage conductance previously presented. It will be recalled that these values of leakage are derived on the basis of 40 pairs of insulators per mile, and are intended to represent, not average values, but the highest values ordinarily obtained under conditions of dry and wet weather. Systems are ordinarily engineered on the basis of the extreme wet weather attenuation values. When a line runs through the more arid parts of the country, however, advantage is

[10] See "Decibel—The Name for the Transmission Unit," by W. H. Martin, *Bell System Tech. Jl.*, January 1929, pp. 1–2.

often taken of this fact by making the repeater spans longer than normal.

A comparison of the attenuation values for a 165-mil open-wire pair when equipped with different types of insulators is presented in Fig. 11. For purposes of comparison with the normal values a curve of the attenuation due to resistance only, representing the ideal condition of zero leakage conductance, also is given in Fig. 11.

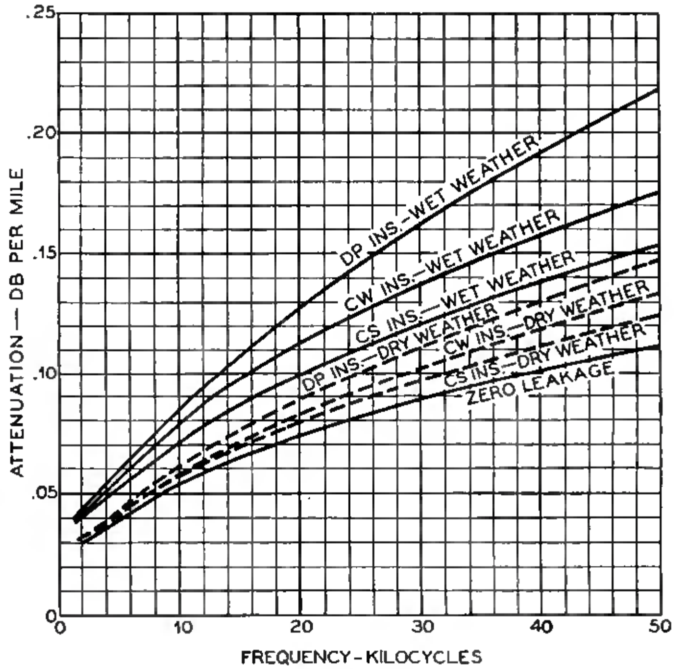


Fig. 11—Attenuation of 165-mil open-wire pair for various conditions of insulation.

The values of line capacitance employed in determining the attenuation of values of Figs. 9, 10, and 11 include an allowance for the average capacitance increase due to the insulators. The attenuation curves shown are strictly applicable to pairs having a wire separation of 12 in., but they are approximately correct for spacings of 8 and 18.25 in.

When the number of pairs of insulators per mile differs greatly from the standard value of 40, a correction is applied to the attenuation values. Special curves make it possible to obtain this correction conveniently. Curves are also available for correcting the standard attenuation values to take care of changes in temperature.

It should be understood that the attenuation of an open-wire pair varies from time to time over a wide range of values, and therefore it is not to be expected that the values of attenuation measured at any particular time will necessarily coincide with the theoretical values. It should also be understood that the attenuation measured on an actual pair never bears the perfectly smooth relation to frequency which is shown on the standard attenuation curves, but exhibits irregularities varying in magnitude according to the irregularities existing on the line. Thus the curve of attenuation as measured on a very well transposed open-wire pair,[11] which is delineated in Fig. 12, represents about as

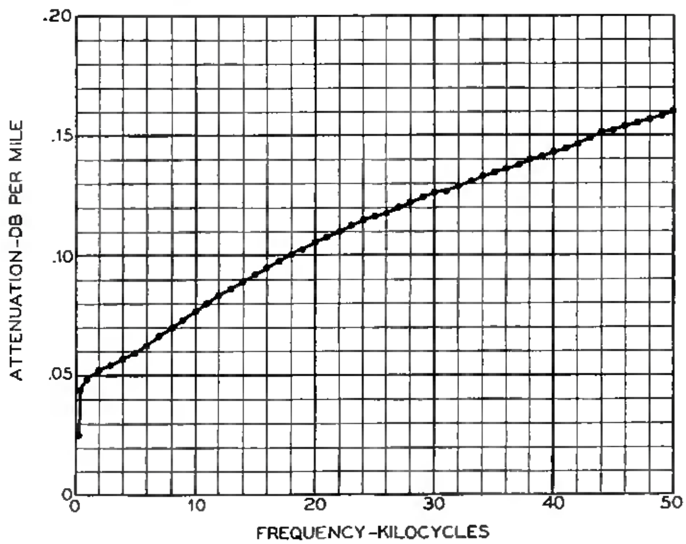


Fig. 12—Attenuation measured on a well-transposed 128-mil open-wire pair with 8-inch spacing.

smooth an attenuation curve as it is possible to obtain on an open-wire circuit. The attenuation values shown on this curve are somewhat lower than the standard values for similar pairs. This is doubtless explained by the fact that the insulators on this particular pair were new, and the further fact that the measurements were made in winter when the temperature was low.

An illustration of the significance of the attenuation data given above may be of interest. One of the longest carrier telephone systems now in service extends from Davenport, Iowa to Sacramento, California, a

[11] Methods of measuring the attenuation, impedance, and crosstalk are discussed in "High-Frequency Measurements of Communication Lines," by H. A. Affel and J. T. O'Leary, *A. I. E. E. Trans.*, Vol. 46, 1927, pp. 504–513.

total distance of about 2100 miles. The highest frequency employed in this system is approximately 28,000 cycles. Using the attenuation values for a 165-mil pair at 28,000 cycles, it appears that the dry weather attenuation for the entire length of this system might be approximately 220 db or less, and the wet weather attenuation about 330 db. This means that without amplification along the line the ratio of the transmitted power to the received power might vary from 10^{22} to 10^{33}. Since the attenuation of a repeater section is ordinarily limited to from 25 to 40 db, ten repeaters are employed to span the total distance, and in order to compensate for the attenuation variations a gain regulating mechanism known as a pilot channel must be used.

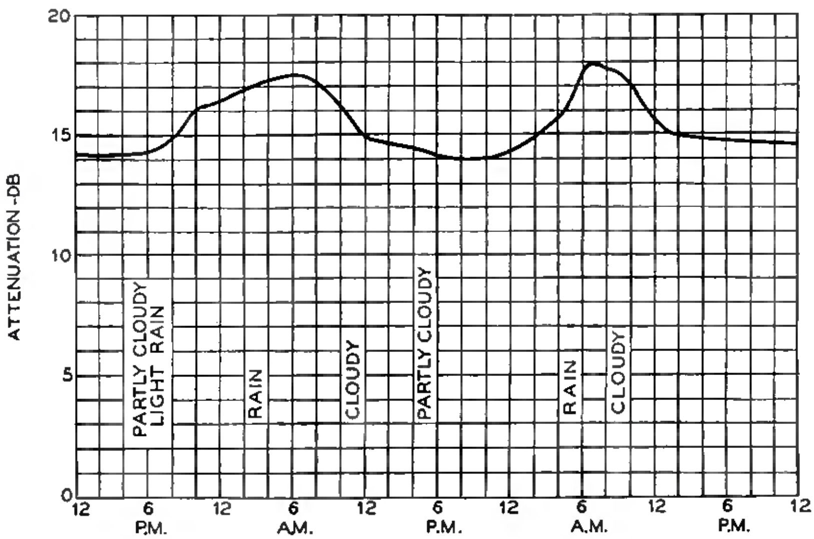


Fig. 13—Measured variations in the attenuation of an open-wire pair.

In the preceding illustration it was assumed that the range of variation in attenuation increases in direct proportion to the length of the circuit. Although this may theoretically be possible, it has been found in practise that the attenuation variations during any given period of time increase less rapidly than the circuit length. The reason for this is that augmenting the length of the circuit obviously reduces the likelihood of experiencing extreme wet weather conditions simultaneously over the entire line. A practical example of how the open-wire attenuation varies from time to time is afforded by Fig. 13, which shows the measured attenuation changes on a line 110 miles long during the period of two light rainstorms.

One further point is of interest in connection with the subject of open-wire attenuation. Inductive or conductive coupling between a pair and the other circuits on the line may result in the absorption of energy in these circuits. Fortunately, the losses due to this cause are small on well transposed lines. On inadequately transposed lines, however, this interaction with other circuits, in addition to producing small losses over a wide range of frequencies, may cause incredibly large losses over a narrow band of frequencies, producing what is known as an "absorption peak" in the attenuation curve. This interesting

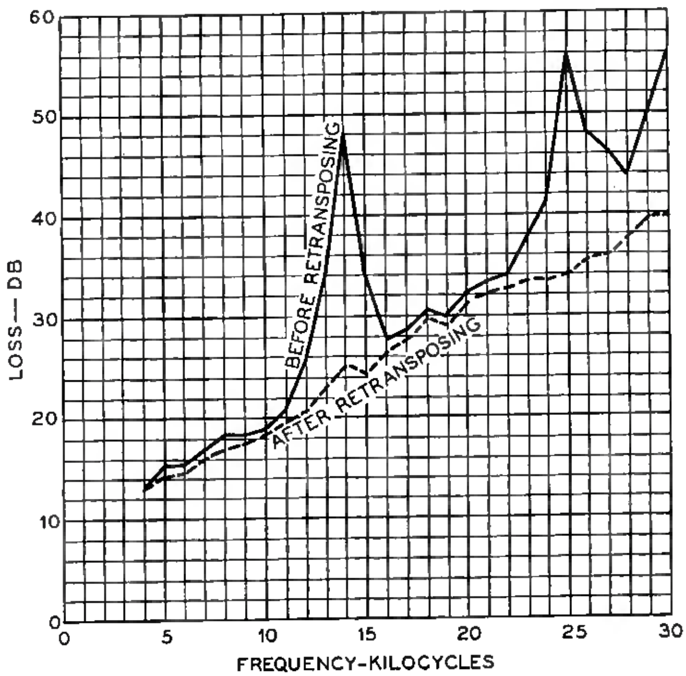


Fig. 14—Absorption peaks on an open-wire pair.

phenomenon is illustrated in the attenuation curves of Fig. 14, which show how two very pronounced absorption peaks on a line about 300 miles in length were smoothed out by the application of improved transpositions. The magnitude of one of these absorption peaks will be appreciated when it is realized that the received power at the peak frequency is about one two-hundredth of that at the adjacent frequencies.

IMPEDANCE:

The characteristic impedance is defined by the well-known formula:

$$Z_0 = \sqrt{\frac{R + jL\omega}{G + jC\omega}} \text{ ohms.} \tag{8}$$

It is doubtless unnecessary to explain why this impedance must be matched in the apparatus.

When R is small compared to $L\omega$ and G is small compared to $C\omega$, the value of Z_0 evidently becomes

$$Z_0 \doteqdot \sqrt{\frac{L}{C}}\cdot \tag{9}$$

This is known as the nominal characteristic impedance. It will be noted that this impedance is a pure resistance. In the carrier range the actual impedance of an open-wire pair is substantially equal to the nominal characteristic impedance.

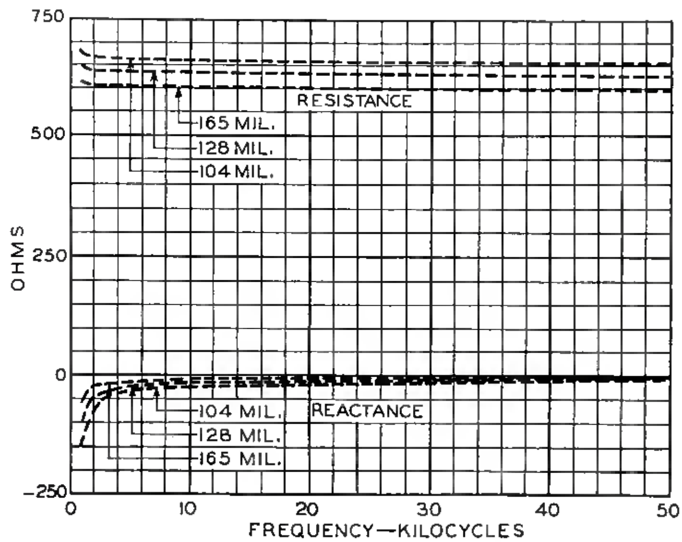


Fig. 15—Carrier-frequency impedance of 12-in. open-wire pairs.

Values of the characteristic impedance in dry weather of open-wire pairs with 12-inch wire spacing are presented in Fig. 15. These values have been derived from the standard values of inductance, capacitance, and leakage conductance and from resistance values at 20 deg. cent. The basis for the impedance value of 600 ohms resistance, which has become almost a tradition in so many phases of telephone work, will be obvious from Fig. 15. The impedance curves for pairs with 8-inch and 18.25-inch spacing are similar to those of Fig. 15, the resistance values for these spacings being about 45 ohms lower, and 40 ohms higher, respectively, than the values shown for 12-inch spacing.

Changes of resistance and leakage conductance due to changing weather conditions have very little effect on the characteristic impedance at frequencies above 1000 cycles. Changes of insulator capacitance due to changing weather conditions or the use of different numbers or types of insulators have an appreciable effect on the impedance. Deviations from the normal spacings between wires which result from the use of drop bracket transpositions also have an important effect upon the impedance.

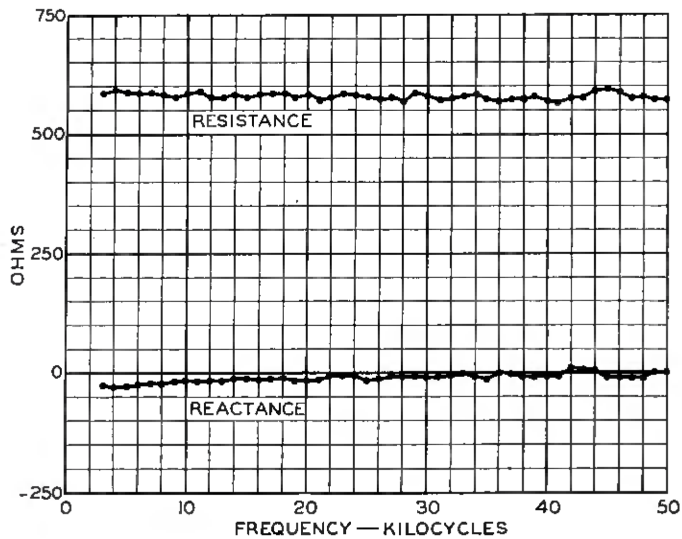


Fig. 16—Impedance measured on a well-transposed 128-mil open-wire pair with 8-in. spacing.

Like the measured attenuation, the impedance which is measured for an open-wire pair is affected by the presence of line irregularities. Hence, the measured impedance is never a smooth function of frequency, but displays slight irregularities throughout the entire range. This is apparent from Fig. 16, which gives a curve of the impedance measured on a well transposed pair. This curve is in remarkably close accord with the generalized values of impedance, the maximum deviation from the theoretical curve being about 2 per cent.

Like the attenuation, the impedance of an open-wire circuit in a narrow band of frequencies may be radically changed by interaction with adjacent circuits. These large irregularities in the impedance commonly accompany absorption peaks in the attenuation, and are, of course, due to the inadequacy of the line transpositions.

Phase Change and Velocity of Propagation

The imaginary component of β of the propagation constant is known as the phase constant because it indicates the change in the phase of the voltage and current in circular radians per unit length of line.

The value of the phase constant is given by

$$\beta^2 = 1/2\ [\sqrt{(R^2 + L^2\omega^2)(G^2 + C^2\omega^2)} + (LC\omega^2 - RG)]. \qquad (10)$$

If R^2 and G^2 are small compared to $L^2\omega^2$ and $C^2\omega^2$ it is clear that

$$\beta \doteqdot \omega\sqrt{LC} \text{ radians per mile.} \qquad (11)$$

For an open-wire pair the value of β is approximately 0.035 radian per mile, or 2 deg. per kilocycle per mile. The latter figure is a convenient one to remember.

The constant β also enters into the familiar expression for the velocity of propagation

$$V = \frac{\omega}{\beta} \text{ miles per second.} \qquad (12)$$

The velocity of propagation on open-wire lines approaches the velocity of light, which is 186,000 miles per second. The velocity is reduced below this value by the increase of capacity due to the presence of the other wires and the insulators, by the internal inductance, and by the presence of resistance and leakage conductance. Values of the velocity of transmission for open-wire pairs are presented in Fig. 17. At frequencies above a few hundred cycles the velocity of transmission is, apart from the effect of line irregularities, practically constant throughout the frequency range.

In the last few years increasing attention has been focused upon the phase characteristic of the open-wire circuit. One of the reasons for this is that different velocities of transmission for different frequency components in a signaling band (which are obtained when the phase shift of the circuit is not a linear function of the frequency) may give rise to what is known as phase distortion, and it may be necessary to correct this distortion by suitable networks.[12, 13]

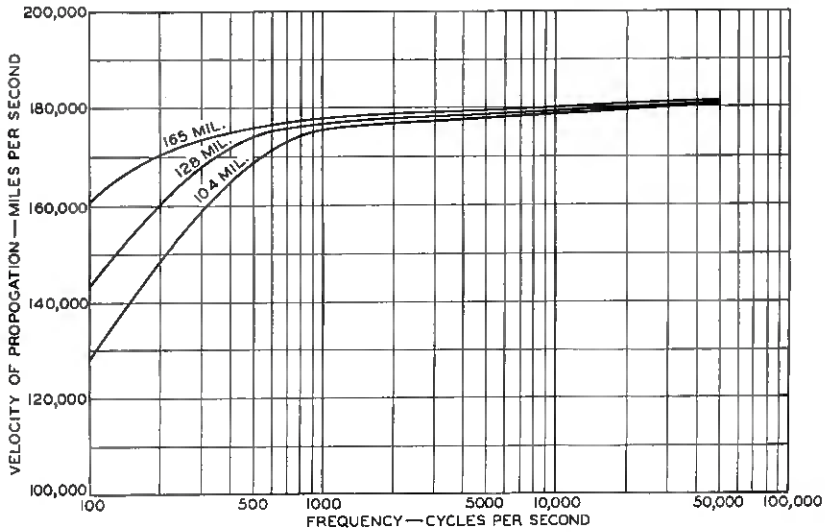


Fig. 17—Velocity of propagation for open-wire pairs.

Characteristics of Phantom Circuits

Phantom circuits, which are derived from two pairs or side circuits by transmitting over the wires of one pair in parallel and using the wires of the other pair in parallel as a return, have been employed in the telephone plant for voice-frequency transmission for a number of years. Their use for carrier transmission is limited chiefly by the difficulty of reducing the cross induction with other circuits at high frequencies.

Phantom circuits are generally derived either from horizontally adjacent non-pole pairs or from vertically adjacent pole pairs. Thus wires 1 and 2 are "phantomed" with wires 3 and 4, wires 5 and 6 with wires 15 and 16, etc. In some of the newer transposition arrangements the non-pole pairs are not phantomed, since it has been found that the omission of the phantom permits the operation of a larger number of carrier systems on one line without excessive mutual interference.

The resistance of a phantom circuit is evidently half of the corresponding value for the side circuit.

[12] See "Distortion Correction in Electrical Circuits with Constant Resistance Recurrent Networks," by O. J. Zobel, *Bell System Tech. Jl.*, July, 1928, pp. 438–534.

[13] See "Wire Transmission System for Television," by D. K. Gannett and E. I. Green, *A. I. E. E. Trans.*, Vol. 46, 1927, pp. 946–953 (*Bell System Tech. Jl.*, October, 1927, pp. 616–632).

If the two wires forming one of the sides of the phantom circuit are designated 1 and 2, and the wires forming the other side are designated 3 and 4, the inductance of the phantom circuit is

$$L = 0.32187 \left[1.1513 \log_{10} \frac{4D_{13}D_{14}D_{23}D_{24}}{D_{12}D_{34}d^2} + \mu\delta \right] \times 10^{-3} \text{ henrys per loop mile} \quad \mathbf{(13)}$$

where D_{13}, D_{23}, etc. represent the spacing between the wires 1 and 3, 2 and 3, etc., and where d, μ, and δ have the same significance as in Equation (**3**).

The capacitance of a phantom circuit in space is

$$C = \frac{0.07766 \times 10^{-6}}{\log_{10} \dfrac{4D_{13}D_{14}D_{23}D_{24}}{D_{12}D_{34}d^2}} \text{ farads per loop mile.} \quad \mathbf{(14)}$$

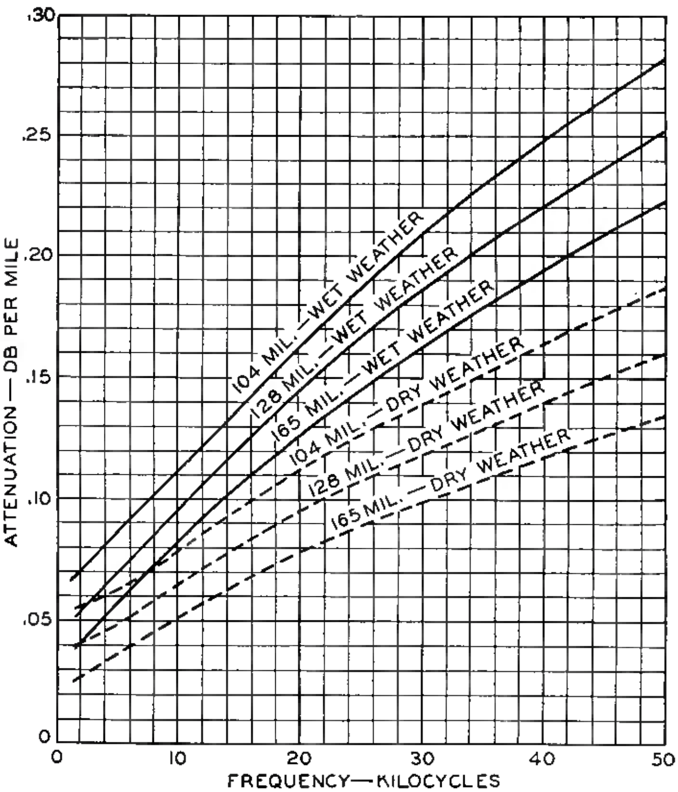


Fig. 18—Attenuation of phantom open-wire circuits.

Equations (**13**) and (**14**) are based on the assumption that the phantom circuit is sufficiently well transposed to secure balanced voltages, currents, and charges on the four wires. The capacitance of the phantom circuit is, of course, affected by the capacitances of its component wires to the other wires of the lead and to ground and also by the capacitances of the insulators.

The losses which contribute to the leakage conductance of phantom circuits are similar to those which have been discussed for side circuits. The ratio of the phantom circuit leakage conductance to that of the side circuit will depend upon the relative magnitudes of these different losses. Varying relations between the different types of losses give a ratio of phantom circuit leakage conductance to side circuit leakage conductance which might conceivably range between the extremes of 1 and 6. For practical engineering purposes, however, the leakage conductance of the phantom circuit is generally assumed to be twice that of the side circuit.

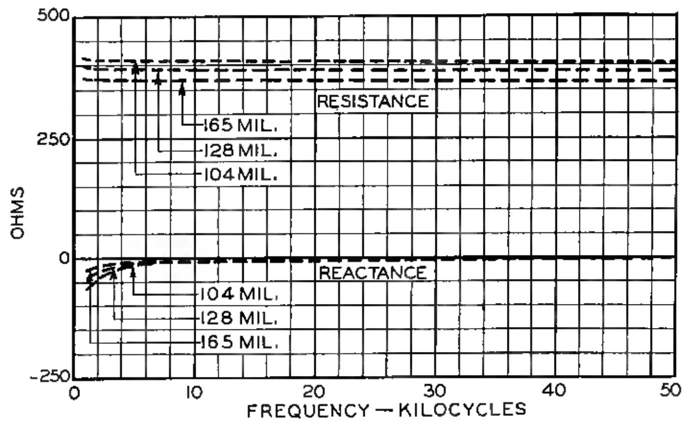


Fig. 19—Impedance of phantom open-wire circuits.

Curves of the attenuation of non-pole pair 12-inch phantom circuits of different gages are depicted in Fig. 18. In most cases the attenuation of the phantom circuit is somewhat less than that of the corresponding side circuit. Frequently, however, the advantage of lower attenuation is under practical conditions more than offset by the large noise and crosstalk effects experienced on the phantom circuits.

Values of the dry weather impedance of 12-inch phantom circuits are presented in Fig. 19. It will be observed that the impedance of a phantom circuit averages about 60 per cent of the impedance of a 12-inch side circuit.

Characteristics of Iron-Wire Circuits

There now exists on the toll lines of the Bell System a small amount of galvanized iron or steel wire. Nos. 12 and 14 B. W. G., with diameters of 109 and 83 mils, respectively, are the gages most commonly found. Steel wire and BB iron wire are both used, the latter being of more frequent occurrence. These iron-wire pairs display such large values of modulation and high-frequency attenuation that they are generally quite unsuitable for carrier transmission.

The skin effect resistance of iron wire may be computed by standard theory, provided that the values of the resistivity ρ and permeability μ are known. The values of ρ for BB iron and steel wire are about 12.8

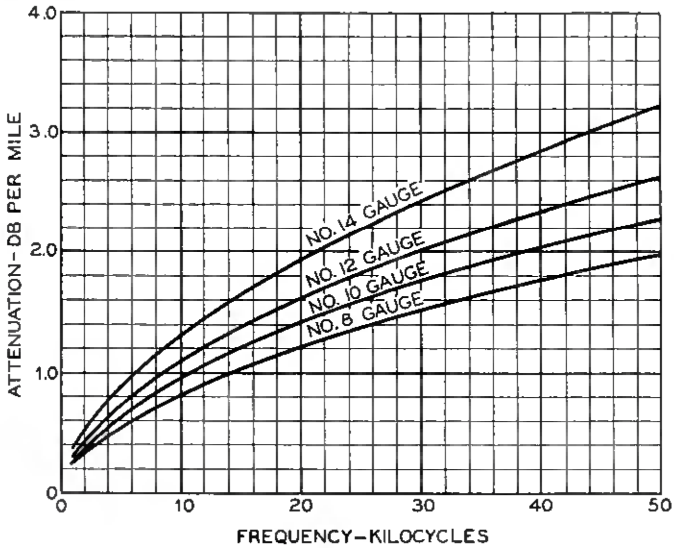


Fig. 20—Attenuation of iron-wire circuits.

and 14.8 microhm-centimeters respectively. It will be noted that these values are about seven and eight times the resistivity of copper. For currents of telephonic magnitude at frequencies from about 500 to 50,000 cycles, it has been found that the value of the permeability μ of iron wire of the above sizes ranges from about 110 to 50. Study of the available data indicates that the best average value of μ at 1000 cycles is about 85 and that the effective value decreases with increasing frequency. Because of the high d.-c. resistance and the large skin effect ratio which results from the high permeability, the resistance of iron wire for alternating currents is extremely great.

Except at low frequencies, where the internal inductance of the iron wire is large, the total inductance of an iron-wire circuit is not far different from that of the corresponding copper circuit. The capacitance and leakage conductance of iron-wire circuits are substantially the same as for similar pairs of copper wire.

Typical attenuation curves for several gages of BB iron wire in dry weather are shown in Fig. 20. These curves are based upon experimental results. It will be noted that the attenuation of an iron-wire circuit averages about ten times that of the corresponding copper-wire circuit. It should be understood that the attenuation values for an iron-wire circuit are subject to rather wide variations in practise, particularly because of the effects of corrosion. There is some change in the attenuation of an iron-wire circuit with change in weather conditions, but this is a relatively small percentage. The impedance of an iron-wire circuit has the same order of magnitude as the impedance of a similar copper-wire circuit.

Appendix I

Temperature Coefficient of Resistance at High Frequencies

In the skin effect literature cited in the text, it is shown that at high frequencies the a.-c. resistance R is

$$R = 0.00979\sqrt{fR_0}, \tag{15}$$

where R_0 is given by Equation (6) and f is the frequency. Differentiating with respect to t, and substituting, we find that

$$A_1 = \frac{1}{R_1}\frac{dR}{dt} = \frac{1}{2}\alpha_1\sqrt{\frac{R_{01}}{R_0}} \tag{16}$$

and for ordinary ranges of temperature

$$A_1 = \frac{\alpha_1}{2}\,. \tag{17}$$

Appendix II

Condition for Minimum Attenuation

The attenuation at high frequencies is given by Equation (7). Differentiating this with respect to $\sqrt{\frac{C}{L}}$, and setting the result equal to zero, it is found that the condition for minimum attenuation is

$$\frac{L}{C} = \frac{R}{G}, \tag{18}$$

which is another way of saying that

$$\frac{R}{2}\sqrt{\frac{C}{L}} = \frac{G}{2}\sqrt{\frac{L}{C}}, \tag{19}$$

so that the resistance and leakage components must be equal in order to have minimum total attenuation.

Transients in Parallel Grounded Circuits, One of Which is of Infinite Length

By LISS C. PETERSON

This paper deals with a mathematical discussion of induction due to transient currents of the forms $I = \sin \omega t$ and $I = e^{-\beta t}$. Formulas and curves are developed for the calculation of the induced voltage in exposed telephone lines due to currents of the above types.

PART I

THE problem of mutual impedance between grounded circuits of infinite length for steady state sinusoidal currents has been treated by a number of authors, and the solution of this problem is now well established.[1, 2, 3] In addition to the steady state voltages induced the transient voltages are also of importance. Rüdenberg [4] and Ollendorf [5] have given approximate solutions for transient voltages due either to d.-c. switching or the sudden flow of a sinusoidal current on the assumption of circular symmetry and for circuits one of which is of infinite length. Since the assumption of circular symmetry holds only for a limited set of conditions it is desirable to develop formulas for the transient induced voltages based on the exact solution for steady state conditions referred to above.

The discussion in this paper will be limited to the case of parallel wires, one of which is of infinite length, and both located on the surface of the earth but insulated from it except at their ends. Disturbing currents of the forms $I = \sin \omega t$ and $I = e^{-\beta t}$ will be assumed.

A more general case with both wires above the earth's surface leads to complicated expressions for the induced voltage not well adapted for engineering use. The restriction to wires on the earth's surface results in appreciable simplification and does not introduce a serious departure from actual conditions.

With these assumptions, the following formulas holding for small and large values of time, determine the induced voltage per unit length on a secondary wire 2 due to the sudden flow of a current $I(t) = \sin \omega t$ in a primary wire 1 infinite in length, separated from wire 2 by a distance x centimeters.

[1] Pollaczek, F., *E. N. T.*, Vol. 3, 1926.
[2] Carson, J. R., *Bell System Technical Journal*, Vol. 5, 1926.
[3] Haberland G., *Z. ang. Math. U. Mech.*, Vol. 6, No. 5, 1926.
[4] Wiss. Veroff. a. d. Siemens-Konzern, Vol. 5, No. 3, 1927.
[5] E. N. T., Vol. 5, No. 3, 1928.

$$V_{12}(t) = \frac{\sin \omega t}{\pi\lambda x^2} - \frac{\omega}{\pi\lambda x^2} e^{-\pi\lambda x^2/t} \left[\frac{t^2}{\pi\lambda x^2} - \frac{2t^3}{(\pi\lambda x^2)^2} + \frac{6t^4 - \omega^2 t^6}{(\pi\lambda x^2)^3} - \frac{24t^5 - 12\omega^2 t^7}{(\pi\lambda x^2)^4} + \cdots \right]$$

and

$$V_{12}(t) = \frac{\sin \omega t}{\pi\lambda x^2} + \frac{2\sqrt{\omega}}{x\sqrt{\pi\lambda}} [\cos \omega t \ \mathrm{kei}'(2x\sqrt{\pi\lambda\omega}) + \sin \omega t \ \mathrm{ker}'(2x\sqrt{\pi\lambda\omega})] - e^{-\pi\lambda x^2/t} \left[\frac{1}{\omega t^2} - \frac{1}{\omega^2}\left(\frac{6}{t^4} - \frac{6\pi\lambda x^2}{t^5} + \frac{(\pi\lambda x^2)^2}{t^6} \right) + \cdots \right]$$

and for such values of time where neither of these series expansions would give very accurate results the following formulas may be used

$$V_{12}(t) = \frac{\sin \omega t}{\pi\lambda x^2} - \frac{\sqrt{A^2 + B^2}}{\pi\lambda x^2} \cos(\omega t - \varphi),$$

$$\tan \varphi = \frac{B}{A} \cdot$$

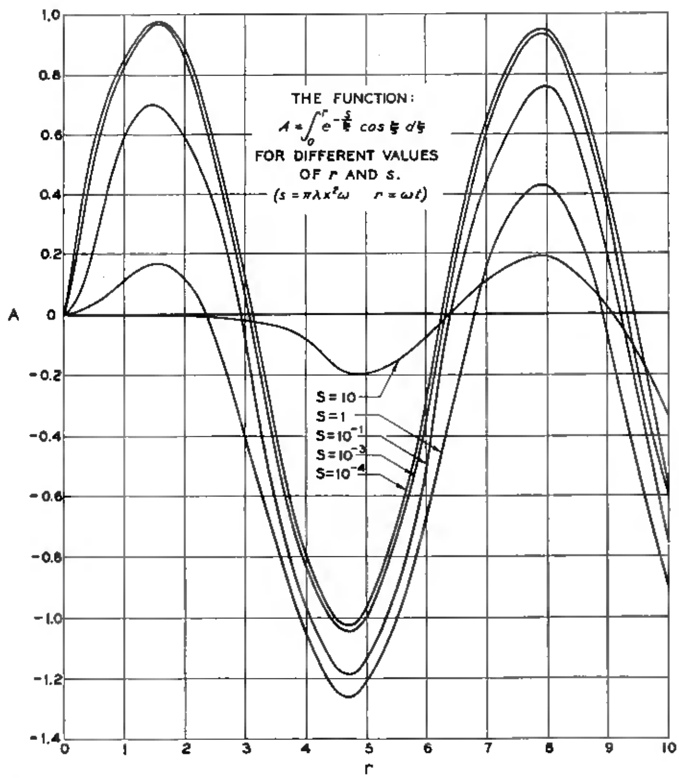


Fig. 1—Plot of the integral A as a function of r for different values of s.

A and B are given by

$$A = \int_0^{\omega t} e^{-\pi\lambda x^2 \omega/\xi} \cos \xi d\xi,$$

$$B = \int_0^{\omega t} e^{-\pi\lambda x^2 \omega/\xi} \sin \xi d\xi.$$

With a disturbing current $I = e^{-\beta t}$ in wire 1, the induced voltage

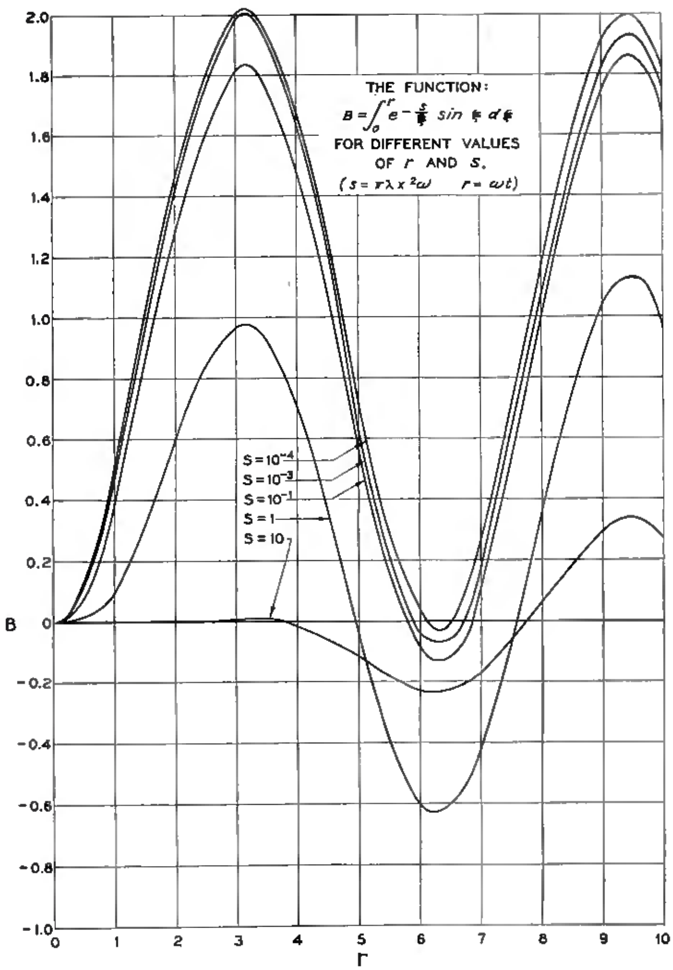


Fig. 2—Plot of the integral B as a function of r for different values of s.

per unit length in wire 2 for small values of time is given by

$$V_{12}(t) = \frac{1}{\pi\lambda x^2}(e^{-\beta t} - e^{-\pi\lambda x^2/t}) + \frac{\beta}{\pi\lambda x^2}e^{-\pi\lambda x^2/t}\left[\frac{t^2}{\pi\lambda x^2} - \frac{2t^3+\beta t^4}{(\pi\lambda x^2)^2} + \frac{6t^4+6\beta t^5+\beta^2 t^6}{(\pi\lambda x^2)^3} - \frac{24t^5+36\beta t^6+12\beta^2 t^7+\beta^3 t^8}{(\pi\lambda x^2)^4} + \cdots\right]$$

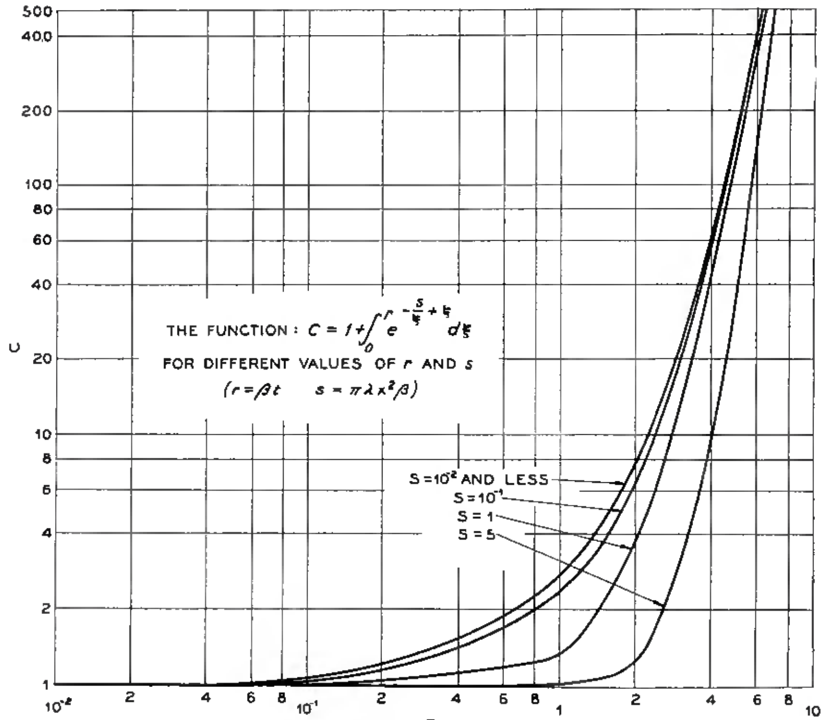


Fig. 3—Plot of the quantity C as a function of r for different values of s.

and for values of time such that the above series can not be used by

$$V_{12}(t) = \frac{1}{\pi\lambda x^2}(Ce^{-\beta t} - e^{-\pi\lambda x^2/t}),$$

where

$$C = 1 + \int_0^{\beta t} e^{-(\pi\lambda x^2\beta/\xi)+\xi} d\xi.$$

Finally, the induced voltage $Z_{12}(t)$ due to a unit step current in wire

1 is determined by

$$Z_{12}(t) = \frac{1}{\pi\lambda x^2}(1 - e^{-\pi\lambda x^2/t}).$$

The functions A, B, and C are plotted in Figs. 1, 2, and 3 for some values of the parameters often to be found in practice.

In these formulas λ is the ground conductivity in electromagnetic c.g.s. units, x the separation between wires in centimeters, t the time in seconds, and $j = \sqrt{-1}$. The functions ker′ and kei′ are related to the Bessel function of the second kind for imaginary arguments defined by G. N. Watson, "Bessel Functions" as follows

$$\text{ker}'(z) \pm j\,\text{kei}'(z) = -j^{\pm 1/2}K_1(zj^{\pm 1/2})$$

Values of these functions are tabulated in Table I of "Bessel Functions for A-C Problems" by H. B. Dwight A. I. E. E. Trans. 1929 pp. 812–820.

The induced voltage is in units of abvolts per cm. which is transformed to volts per mile by the factor 1.61×10^{-4}.

Part II

The second part of this paper will be devoted to a discussion of the theory leading to the above results.

Consider a system of two wires, 1 and 2, wire 1 being of infinite length, parallel with each other, with the heights h_1 and h_2 above earth and separated by a distance x. The general problem is to calculate the voltage on wire 2 as a function of time due to the sudden flow of a current in wire 1, this current being zero before $t = 0$ and $I(t)$ thereafter. Let the voltage on wire 2 due to a unit current step, that is, a current equal to zero before $t = 0$ and unity after $t = 0$, be denoted by $Z_{12}(t)$, then the voltage due to a current $I(t)$ is given by

$$V_{12}(t) = \frac{d}{dt}\int_0^t Z_{12}(\tau)I(t - \tau)d\tau.\text{[6]} \tag{1}$$

The fundamental quantity thus necessary in the solution of the problem is $Z_{12}(t)$. This quantity completely determines the voltage $V_{12}(t)$ for all types of disturbing currents. $Z_{12}(t)$ may be written as a Fourier integral:

$$Z_{12}(t) = \frac{1}{2\pi}\int_{-\infty}^{+\infty} \frac{e^{j\omega t}}{j\omega} Z_{12}(\omega)d\omega, \tag{2}$$

where $Z_{12}(\omega)$ is the mutual impedance for periodic earth currents and

[6] Carson, Electric Circuit Theory and the Operational Calculus, page 16.

is given by [7]

$$Z_{12}(\omega) = 2j\omega \log \frac{\rho''}{\rho'} + 4\omega \int_0^\infty [\sqrt{\mu^2 + j} - \mu] e^{-(h_1+h_2)\mu\sqrt{\alpha}} \cos \mu x \sqrt{\alpha}\, d\mu. \quad (3)$$

where

$$\alpha = 4\pi\lambda\omega$$
$$\rho'' = \sqrt{(h_1 + h_2)^2 + x^2}$$
$$\rho' = \sqrt{(h_1 - h_2)^2 + x^2}$$

$Z_{12}(\omega)$ is limited by the neglect of displacement currents to frequencies such that the propagation constant is a small quantity in c.g.s. units. To obtain $Z_{12}(t)$ it is necessary to integrate over all frequencies as shown by (2); this introduces an approximation in all results for small values of the time.

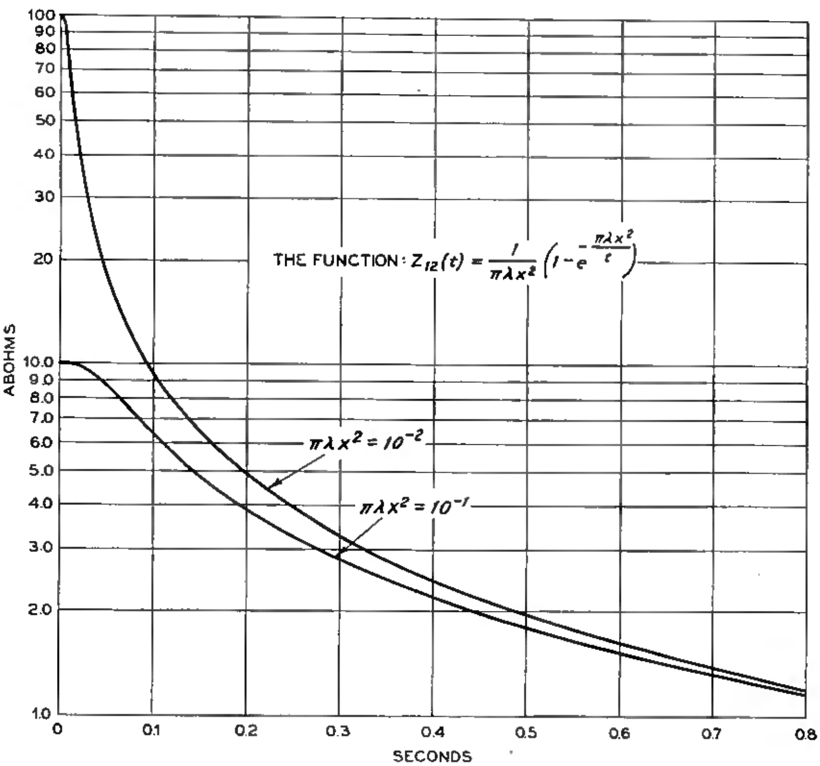


Fig. 4—Plot of the voltage $Z_{12}(t)$ on wire 2 as a function of t for two different values of separation or conductivity as given by the product $\pi\lambda x^2$.

[7] *Bell System Technical Journal,* Vol. V, page 544, October, 1926.

I am indebted to Dr. F. H. Murray of the American Telephone and Telegraph Company for the following solution of $Z_{12}(t)$ as given by (2):

$$Z_{12}(t) = \frac{2h}{|M|^2\sqrt{\lambda}\sqrt{t}} - \left[\frac{1}{2\lambda}\left(\frac{1}{M_1^2}+\frac{1}{M_2^2}\right)\right] + \frac{1}{2\lambda}\left[\frac{e^{\frac{M_1^2\lambda}{t}}}{M_1^2}\ \mathrm{erfc}\,\frac{M_1\sqrt{\lambda}}{\sqrt{t}} + \frac{e^{\frac{M_2^2\lambda}{t}}}{M_2^2}\ \mathrm{erfc}\,\frac{M_2\sqrt{\lambda}}{\sqrt{t}}\right], \quad (4)$$

where

$$|M| = |M_1| = |M_2|,$$
$$M_1 = (h - jx)\sqrt{\pi},$$
$$M_2 = (h + jx)\sqrt{\pi},$$
$$h = h_1 + h_2$$

and

$$\mathrm{erfc}\ Z = 1 - \mathrm{erf}\ Z = 1 - \frac{2}{\sqrt{\pi}}\int_0^Z e^{-x^2}dx.$$

Taking the limit of equation (4) as h approaches zero there results

$$Z_{12}(t) = \frac{1}{\pi\lambda x^2}(1 - e^{-\pi\lambda x^2/t}), \quad (5)$$

which formula is of fundamental importance in the present analysis.[8] This equation is also plotted on Fig. 4 for two different values of $\pi\lambda x^2$.

Assuming now $I(t) = \sin \omega t$ formula (1) gives

$$V_{12}(t) = \frac{\sin \omega t}{\alpha} - \frac{\omega}{2\alpha}\left[e^{\beta t}\int_0^t e^{-(\alpha/\tau)-\beta\tau}d\tau + e^{-\beta t}\int_0^t e^{-(\alpha/\tau)+\beta\tau}d\tau\right], \quad (6)$$

where

$$\alpha = \pi\lambda x^2,$$
$$\beta = j\omega.$$

[8] This formula can readily be checked in the following manner. The mutual impedance between wires on the surface of the ground is

$$Z_{12}(\omega) = \frac{1}{\pi\lambda x^2} - \frac{\gamma}{\pi\lambda x}K_1(\gamma x),$$

where $\gamma = \sqrt{4\pi\lambda j\omega}$ and K_1 is the Bessel function of the second kind with imaginary argument defined by Watson, "Bessel Functions."

Replace $j\omega$ by p, and interpret the function of p so obtained according to operational methods. The first term is independent of p and therefore of t. The second term is transformed according to the equivalent

$$\alpha\sqrt{p}K_1(\alpha\sqrt{p}) = e^{-\alpha^2/4t},$$

given as pair 922 in G. A. Campbell's paper "The Practical Application of the Fourier Integral," *Bell System Technical Journal*, Oct. 1928.

Equation (5) is then immediately obtained.

The integrals appearing in (6) are apparently not known in closed form. Series expansions holding for small and large values of time may be derived however.

By successive integration by parts we obtain:

$$\int_0^t e^{-(\alpha/\tau)+\beta\tau}d\tau = e^{-(\alpha/t)+\beta t}\left[\frac{t^2}{\alpha} - \frac{2t^3+\beta t^4}{\alpha^2} + \frac{6t^4+6\beta t^5+\beta^2t^6}{\alpha^3} - \frac{24t^5+36\beta t^6+12\beta^2t^7+\beta^3t^8}{\alpha^4}+\cdots\right]. \quad (7)$$

$e^{\beta t}$ appearing on the right hand side of equation (7) is cancelled by $e^{-\beta t}$ appearing before the integral, and similarly for the first term in brackets in equation (6). In the complete expression for the voltage odd powers of β cancel and we have:

$$V_{12}(t) = \frac{\sin\,\omega t}{\pi\lambda x^2} - \frac{\omega}{\pi\lambda x^2}e^{-\pi\lambda x^2/t}\left[\frac{t^2}{\pi\lambda x^2} - \frac{2t^3}{(\pi\lambda x^2)^2} + \frac{6t^4-\omega^2t^6}{(\pi\lambda x^2)^3} - \frac{24t^5-12\omega^2t^7}{(\pi\lambda x^2)^4}+\cdots\right]. \quad (8)$$

For large values of time equation (6) is written as

$$V_{12}(t) = \frac{\sin\,\omega t}{\alpha} + \frac{\omega}{2\beta}\left[e^{-\beta t}\int_0^\infty \frac{e^{-(\alpha/\tau)+\beta\tau}}{\tau^2}d\tau - e^{\beta t}\int_0^\infty \frac{e^{-(\alpha/\tau)-\beta\tau}}{\tau^2}d\tau\right] + \frac{\omega}{2\beta}\left[e^{\beta t}\int_t^\infty \frac{e^{-(\alpha/\tau)-\beta\tau}}{\tau^2}d\tau - e^{-\beta t}\int_t^\infty \frac{e^{-(\alpha/\tau)+\beta\tau}}{\tau^2}d\tau\right], \quad (9)$$

where the integrals between zero and infinity correspond to the steady state condition while the integrals between t and infinity give the transient distortion. The integral between t and infinity may be evaluated in a manner quite similar to that used above. The result with plus sign for β is

$$\int_t^\infty \frac{e^{-(\alpha/\tau)+\beta\tau}}{\tau^2}d\tau = e^{-(\alpha/t)+\beta t}\left[\frac{1}{\beta t^2} + \frac{1}{\beta^2}\left(\frac{2}{t^3}-\frac{\alpha}{t^4}\right) + \frac{1}{\beta^3}\left(\frac{6}{t^4}-\frac{6\alpha}{t^5}+\frac{\alpha^2}{t^6}\right) + \frac{1}{\beta^4}\left(\frac{24}{t^5}-\frac{36\alpha}{t^6}+\frac{12\alpha^2}{t^7}-\frac{\alpha^3}{t^8}\right)+\cdots\right]. \quad (10)$$

The integrals between 0 and infinity are evaluated by:

$$\int_0^\infty \frac{e^{-(\alpha/\tau)\pm\beta t}}{\tau^2}d\tau = \frac{2}{\sqrt{\alpha}}(\mp\,\beta)^{1/2}K_1(2\sqrt{\mp\,\alpha\beta}), \quad (11)$$

in which the real and imaginary parts of the right hand side may be expressed by the ker′ and kei′ functions by the relation already given.

The complete expression for the voltage is:

$$V_{12}(t) = \frac{\sin \omega t}{\pi \lambda x^2} + \frac{2\sqrt{\omega}}{x\sqrt{\pi\lambda}}[\cos \omega t \ \mathrm{kei}'(2x\sqrt{\pi\lambda\omega}) + \sin \omega t \ \mathrm{ker}'(2x\sqrt{\pi\lambda\omega})]$$
$$- e^{-\pi\lambda x^2/t}\left[\frac{1}{\omega t^2} - \frac{1}{\omega^2}\left(\frac{6}{t^4} - \frac{6\pi\lambda x^2}{t^5} + \frac{(\pi\lambda x^2)^2}{t^6}\right) + \cdots\right]. \quad (12)$$

For such values of time where neither of the formulas (8) or (12) give very accurate results it is necessary to perform mechanical integration.

In so doing it is convenient to introduce a new variable ξ of integration. Let $\xi = \omega\tau$, and the integrals become

$$\int_0^t e^{-(\alpha/\tau)\pm j\omega\tau} d\tau = \frac{1}{\omega}\int_0^r e^{-(s/\xi)\pm j\xi} d\xi$$
$$= \frac{1}{\omega}\left[\int_0^r e^{-s/\xi} \cos \xi d\xi \pm j \int_0^r e^{-s/\xi} \sin \xi d\xi\right],$$

where

$$\left.\begin{aligned} s &= \alpha\omega = \pi\lambda\omega x^2, \\ r &= \omega t. \end{aligned}\right\} \quad (13)$$

Now let

$$A(s, r) = \int_0^r e^{-s/\xi} \cos \xi d\xi, \quad (14)$$

$$B(s, r) = \int_0^r e^{-s/\xi} \sin \xi d\xi, \quad (15)$$

and formula (6) becomes

$$V_{12}(t) = \frac{\sin \omega t}{\pi\lambda x^2} - \frac{\sqrt{A^2 + B^2}}{\pi\lambda x^2} \cos(\omega t - \varphi), \quad (16)$$

where

$$\tan \varphi = \frac{B}{A}. \quad (17)$$

The values of the functions A and B are given in Figs. 1 and 2 for some values of s and r which are frequently met in practice.

Assuming finally $I(t) = e^{-\beta t}$ equation (1) gives after simplifications

$$V_{12}(t) = \frac{1}{\pi\lambda x^2}(e^{-\beta t} - e^{-\pi\lambda x^2/t}) + \frac{\beta e^{-\beta t}}{\pi\lambda x^2}\int_0^t e^{-(\pi\lambda x^2/\tau)+\beta\tau} d\tau. \quad (18)$$

For small values of t the series expansion (7) may be used. The result for this case then becomes

$$V_{12}(t) = \frac{1}{\pi\lambda x^2}(e^{-\beta t} - e^{-\pi\lambda x^2/t}) + \frac{\beta}{\pi\lambda x^2} e^{-\pi\lambda x^2/t}\left[\frac{t^2}{\pi\lambda x^2} - \frac{2t^3 + \beta t^4}{(\pi\lambda x^2)^2} + \frac{6t^4 + 6\beta t^5 + \beta^2 t^6}{(\pi\lambda x^2)^3} - \frac{24t^5 + 36\beta^2 t^6 + 12\beta^2 t^7 + \beta^3 t^8}{(\pi\lambda x^2)^4} + \cdots\right]. \quad (19)$$

For large values of time, introduce a new variable $\xi = \beta\tau$ of integration in the integral in (18). Then

$$V_{12}(t) = \frac{1}{\pi\lambda x^2}(Ce^{-\beta t} - e^{-\pi\lambda x^2/t}) \quad (20)$$

where

$$\left.\begin{aligned} C &= 1 + \int_0^r e^{-(s/\xi)+\xi} d\xi, \\ r &= \beta t, \\ s &= \pi\lambda x^2\beta. \end{aligned}\right\} \quad (21)$$

The values of C are given on Fig. 3 for important ranges of r and s. For s equal to and less than 10^{-2}, C is for practical purposes independent of s.

I am indebted to Dr. F. H. Murray, and Mr. R. M. Foster of the American Telephone and Telegraph Company for valuable suggestions during the course of this work, and to Miss R. Pedersen who carried out all the numerical calculations.

Impedance Correction of Wave Filters

Development of Impedance Requirements

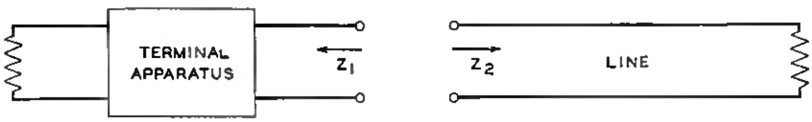
By E. B. PAYNE

The present importance of wave filter impedance correction arises chiefly from its relation to crosstalk in carrier systems. Briefly, it appears that line transpositions, an effective remedy for many types of crosstalk, are less satisfactory when directed against the so-called "reflected near-end crosstalk" and "reflected far-end crosstalk" produced when waves reflected from the junctions between lines and repeater equipment of carrier systems induce currents in neighboring systems. The expense of the elaborate transposition scheme necessary for a substantial reduction in these types of crosstalk makes it desirable to diminish the amplitude of the reflected wave as far as possible by the improvement of the impedance match between lines and repeaters. A detailed study shows that this is most conveniently done by terminating the filters in the repeater by sections whose image impedances at one end match the main body of the filter, while at the other they approximate constant resistances, matching the terminal impedances.

The development of appropriate filter terminating sections has passed through a number of stages. The earliest filters gave reflection coefficients as great as 50% to 60% in the useful transmission band. The invention of "m-derived" and "x-terminated" filters, plus a number of more or less empirical schemes, made it possible to obtain reflection coefficients ranging from 10% to 15% in the useful band. Recent progress has resulted chiefly from the development of a series of sections, the simplest of which is equivalent to the m-derived type, while the others, of progressively increasing complexity, give progressively better approximations to the ideal characteristic. The use of the more complicated sections has made it possible to reduce filter reflection coefficients to the order of 2%, or even less. At present the chief limitation appears to be the difficulty of manufacturing filters with sufficient precision to allow the theoretical characteristics to be realized. The paper is illustrated by figures showing the various stages of this progress as they are exemplified in actual designs.

THE rapid increase in the demand for long distance or toll telephone service in recent years led to the introduction, about 1920, of carrier systems as a means of securing more intensive use from long telephone lines. The growth of these circuits has resulted, still more recently, in the multiplication of the number of carrier systems in use and in the close association of several similar or different carrier systems on a single pole-line. This development raised a number of totally new engineering problems and demanded careful reconsideration of many other questions of comparatively small importance in earlier systems.

Among the factors thus brought into prominence by carrier system development, the chief, for the purpose of this paper, is the impedance mismatch between telephone lines and repeaters or terminal apparatus. The components of a complete transmission system, such as the line itself, various transmission networks, amplifiers, modulators, electro-acoustic apparatus, etc. are quite dissimilar physically and as we might naturally expect, these physical differences manifest themselves in many instances as pronounced dissimilarities in the forms of the impedance-frequency characteristics. For example, the characteristic impedance of a uniform line varies smoothly with frequency, but that of a wave filter changes abruptly as we go from the transmitting to the attenuating range. In spite of the possibility of changing the general impedance level by the insertion of a transformer such inherent "incompatibilities of temperament" between the characteristic impedances of the various components of the telephone circuit must lead normally to impedances which resemble each other only in narrow frequency bands and which may differ widely over large and important portions of the frequency spectrum. In default of some method of extending the range of similarity, most long telephone circuits will exhibit wide impedance mismatches or irregularities at numerous junction points.



REFLECTION COEFFICIENT = $\frac{Z_1 - Z_2}{Z_1 + Z_2}$

Fig. 1—Junction of line and terminal apparatus illustrating impedances which determine the reflection coefficient.

In voice frequency circuits or in carrier circuits which are not in close physical association, impedance irregularities are of importance only insofar as they affect transmission efficiency.* In addition to modifying the current which proceeds onward toward the receiving device, however, an impedance difference at any junction produces a reflected wave which retraverses the circuit toward the sending end. A convenient measure of this second effect is found in the "reflection coefficient" which may be defined as the vector difference of the two impedances looking both ways from any junction divided by their vector sum (see Fig. 1) and is equal both in magnitude and phase to the ratio between the reflected wave and the wave originally propagated.

The effect of reflection of considerable magnitude on transmission is slight. Indeed relatively large reflection may actually improve the transmission characteristic of certain circuits. In voice frequency circuits and in carrier circuits which are not operated over lines in close

* In two wire repeatered circuits reflection causes echoes which are one of the limiting factors of such circuits. These circuits are, however, outside the scope of this paper.

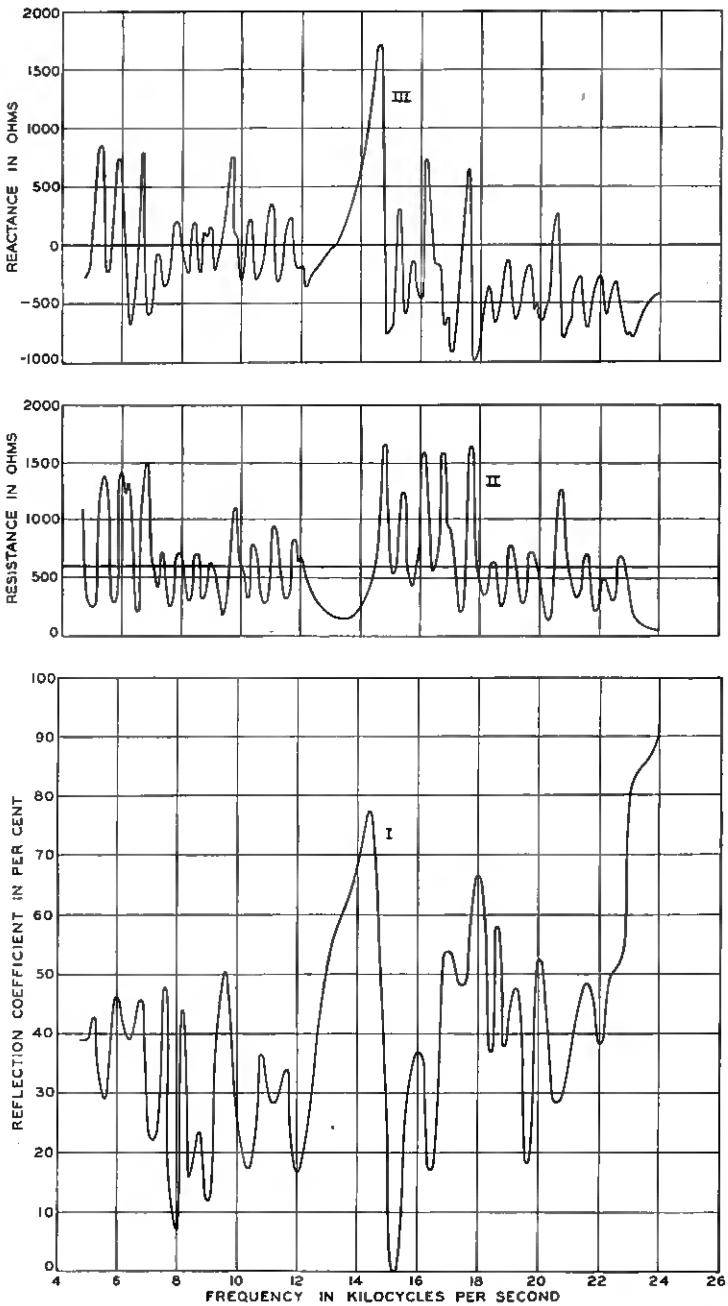


FIG. 2—Impedance and reflection coefficient of an early carrier telephone system.

physical association, transmission is the only consideration. Consequently relatively large reflection coefficients are not objectionable in such circuits. Curve I of Fig. 2 shows the reflection coefficient of an early carrier system which was not intended to work with other systems of the same type. Large as these reflections appear relative to present day standards they did not seriously impair the transmission of the system. It is also true that appreciably better results could not have been obtained with the design technique available when the filters for the system were developed.

As the increase in the demand for long distance traffic made it necessary to associate systems on the same pole line the situation became radically different through the introduction of a new factor crosstalk between systems. Crosstalk between systems at carrier frequencies is inherently large and the methods of reducing it expensive. The reflections due to the mismatching in systems increase the crosstalk between them by introducing a type of interference known as "reflected near-end crosstalk." This type of crosstalk can be made negligible only by making the impedance mismatching in the two systems very small. Since "near-end crosstalk" contributes heavily to the cost of the arrangements for reducing crosstalk between carrier systems the substantial elimination of impedance irregularities between lines and the filters and associated apparatus composing the terminals of systems becomes of great economic importance.

As this need for reducing and ultimately eliminating these irregularities appeared a series of improvements in design technique have been developed, each better than its predecessor, which have culminated in a technique which appears to be adequate for the purpose. It differs in many essential features from the others and leads to a new type of filter section which is not of the standard recurrent ladder type. It is the purpose of this paper to give some idea of the relation of crosstalk to impedance mismatching, show how the successive stages of the filter development have grown out of the system requirements and finally to present an outline of the final technique. The accompanying paper. "A Method of Impedance Correction." by H. W. Bode gives this technique in detail.

Impedance Irregularities and Crosstalk

The ultimate relation between reflection and crosstalk between two lines which are associated with a number of others on telephone poles is extremely complicated. An idea of the principles which underlie the relationship may be obtained by considering only two of the circuits and assuming that the others have been temporarily removed.* If these

* Since these two circuits consist of two pairs of wires, there is a potential phantom

two telephone lines parallel each other, as in Fig. 3, they will be electrically coupled through mutual inductance and capacity. Currents flowing in one will consequently produce crosstalk currents in the other. When the subscriber at the west end of the line (A) is talking, waves initiated by his voice will cross from line (A) to line (B) at adjacent points along the entire length of the lines.

Crosstalk entering line (B) at a typical point may traverse four chief paths. It may (1) flow directly back to the west end, (2) flow onward to B'' and be reflected back to the west end, (3) flow directly onward to the east terminal, and (4) flow backward to B' and thence, by reflection, to B''. There are of course an infinite number of other paths involving multiple reflections but the reduction in amplitude caused by successive

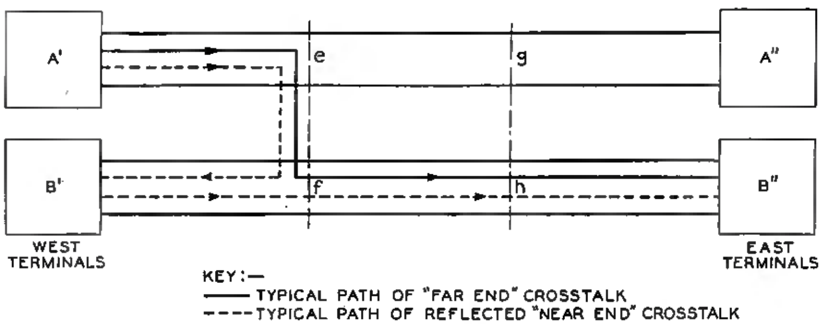


Fig. 3—Diagram illustrating relation between impedance mis-matches and crosstalk in carrier systems.

reflection and line attenuation makes these negligible in comparison with the others. The first two of these four possibilities of crosstalk production can be eliminated immediately. Modern carrier systems are so designed that conversations going in one direction are carried by one band of frequencies and those travelling in the opposite direction by a different band. Currents entering the west terminal, whether they follow a direct path such as $AefB'$, or are first reflected at B'', making the typical path $A'efB''B'$, are therefore eliminated by the filters in the terminal office. Crosstalk currents of the third type ("far end crosstalk") following the typical path $A'efB''$ cannot be eliminated since they fall within the frequency band used by the subscriber at B'' for listening. We may observe, however, that these currents traverse the same length of line in travelling from A' to B'' no matter what the point (such as *ef* in the diagram) at which they cross from one line to the other. Since both lines are alike crosstalk currents will be attenuated and shifted in phase by the attenuation and phase shift of a single

circuit which, as far as crosstalk is concerned, constitutes a third circuit. The effect of this circuit on the crosstalk of the other two is by no means negligible but consideration of it is omitted herein in order to simplify the presentation of other important relations fundamental to crosstalk.

full length line. Moreover the components of this type of crosstalk due to magnetic and capacitative coupling are nearly out of phase and so one tends to neutralize the other. As a consequence of this equal effect of the line characteristic on all the components which reach the receiver at B'' the resultant crosstalk can theoretically be eliminated at all frequencies when only two circuits are present by a single transposition (crossing the wires) in the center of either line.

Crosstalk currents of the fourth type, "reflected near-end crosstalk" following such paths as $A'efB'B''$ and $AghB'B''$, cannot, however, be disposed of so easily. The length of line traversed by the component currents which make up the resultant crosstalk depends upon the point at which they cross from one line to the other, and they will therefore be affected in various fashions by the line attenuation and phase shift. The transposition scheme required to eliminate crosstalk resulting from these currents will consequently depend, at any frequency, upon the length of the line and upon its phase and attenuation characteristics at that frequency. Complete elimination of crosstalk of the fourth type cannot be secured, even for two circuits over a finite frequency band, from a finite number of transpositions.

When other lines are adjacent to the two we have considered the problem of reducing "far-end" and "near-end" crosstalk by transpositions is still more complicated. With a number of lines it is no longer even theoretically possible to eliminate far-end crosstalk by a single transposition. It is, in general true, however, that the cost of a transposition scheme adequate for far-end crosstalk is still much less than that of the elaborate system of transposition required to reduce near-end crosstalk to tolerable values. From an economic standpoint therefore, the cost of transpositions required for near-end crosstalk is usually the main feature to be considered.

Impedance Correction an Economic Means of Controlling Crosstalk

Another method of reducing this near-end crosstalk, and one which experience has shown to be much cheaper than an elaborate transposition scheme is found in the reduction of the reflection coefficient between the line and the repeaters. Obviously the magnitude of the reflected near-end crosstalk depends upon the amount of the impedance mismatch at the junction between the line and the terminal equipment (e.g. at B' in Fig. 3). It can be made as small as we please, even with a very simple transposition scheme, if the reflection coefficient at line-repeater junctions can be sufficiently reduced. No serious mismatches would occur if the impedances of repeaters and terminal equipment were that of the modulators or amplifiers, since at carrier frequencies the impedances of modulators, amplifiers and telephone lines approximate constant resistances. The interposition of filters between lines and modulating or amplifying apparatus, however, normally produces large reflection coefficients. Since the filters in addition to being the apparatus immediately responsible for mismatching, are also inexpensive and easily controlled in comparison with the line, they furnish the most promising field for the reduction of reflection coefficients.

Relation between Actual and Image Impedances

The reflection coefficient which determines the amount of crosstalk exhibited by the system involves directly only the line impedance and the *actual* impedance characteristic of the filter system. In order to understand the peculiarities of the actual impedance of a filter, however, it is necessary to give prior consideration to its *characteristic*, or *image* impedances. The image impedances of any transmitting device are defined as the impedances with which the device must be terminated at both ends if the impedances looking both ways at each pair of terminals are to be matched. In other words, they are the impedances with which the structure must be terminated if no reflections are to occur. Filter sections of different physical configuration and with different attenuation characteristics often have the same image impedance characteristics. Practical filter designs are therefore usually composite structures containing several different types of sections. Internal reflections are avoided by so choosing the arrangement of the section that the image impedance characteristics at all section junctions are matched. Under these circumstances the image impedance characteristics of the complete structure are the same as those of its terminating sections. The image impedance characteristic of typical low pass filter sections is shown in Fig. 4-*A*. A corresponding curve for band pass filters is given in Fig. 4-*B*. The image impedance characteristics are given only for the transmitting bands of the filters since, as previously indicated, the filters themselves suppress crosstalk in the attenuating regions making it unnecessary to control impedances outside the transmitting range. The associated equipment, such as lines and modulators, with which the filters are terminated, are approximately constant pure resistances, and may be represented in the transmitting range by the block type characteristics drawn over the curves.

The relation between the image impedance properties of filters and the actual impedance presented by a repeater or terminal to the line can be understood from the simplified circuit diagram of a typical carrier terminal (Fig. 5). Upon examining the figure we note that there are a number of junctions at which rounded filter image impedance characteristics such as those shown in Fig. 4-*A* or Fig. 4-*B* face the block type image impedances of the same figure. In the system of Fig. 5, for example such junctions occur at *B*, *C*, *D*, *E*, and to some

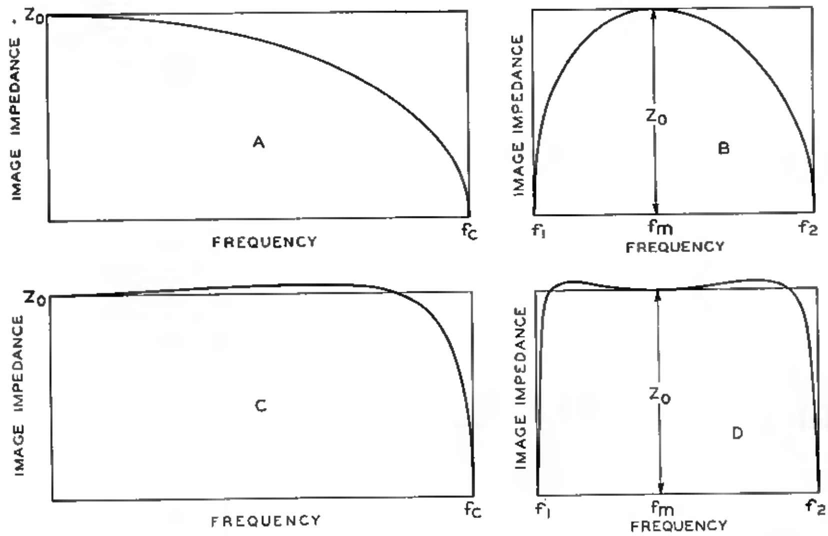


Fig. 4—Image impedances of "constant-*k*" and "m-derived" low-pass and band-pass sections.
Figs. 4-A and 4-B—"constant-*k*" sections.
Figs. 4-C and 4-D—"*m*-derived" sections.

extent at *A*. It is evident, of course, that reflected waves will be produced at these points, and since impedance differences occur at several junctions a wide variety of multiple reflections may exist.

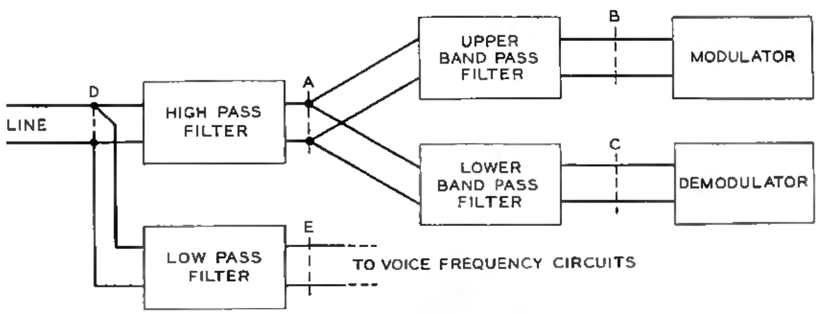


Fig. 5—Simplified circuit diagram of a typical carrier terminal.

Upon reaching the line terminals, all of these reflections combine with the wave originally propagated to determine the actual current entering the structure. The reflected waves are of course diminished in magnitude in traversing the filters intervening between the line terminals and the junctions at which mismatches occur. The attenuation of most filters within transmitting bands is so small however, that the waves may be of appreciable magnitude even after several reflections. The filter phase shift within these frequency bands, on the other hand, is large and varies rapidly with frequency. The reflected waves may therefore combine at the input terminals in almost any fashion, and the effect they produce upon the input current will vary rapidly and violently as we proceed along the frequency scale. With given line impedance and voltage, however, the actual impedance of the terminal is related in simple fashion to the actual current entering it. The impedance, therefore, shows correspondingly wide fluctuations. The extremely irregular impedance and reflection coefficient characteristics of Fig. 2 exemplify the effect of reflections from the further junction points of the system. Another example is furnished by the curve of Fig. 7, which shows the reflection coefficient between the actual impedance of the filter of Fig. 6 when terminated in the line resistance, and that resistance. The humps of the curve come at frequencies whose phase shift is such that the wave reflected from the far end accentuates the departure of the near end image impedance from the desired value. The valleys correspond either to points at which the image impedances are ideal or to values of filter phase shift which cause the reflections at the two ends of the filter to correct one another.

Terminal Impedances Best Corrected by Special Type of Filter Section

Close impedance correction of these complicated characteristics seems hopeless. In order to keep the problem within manageable limits it is necessary to destroy the reflected waves at their source by preventing mismatches at all junctions between filters and other apparatus within the transmitting bands. The technical problem can consequently be reduced to the construction of a new type of filter section for use at terminations, the image impedance of the new section showing at one end a close approximation to the block type terminating impedance characteristic of Fig. 4-*A* or 4-*B* (i.e. a constant resistance) while at the other end it has the conventional rounded filter image impedances also shown on these figures, thus matching the standard sections forming the main bulk of the structure.

Early Improvements

Methods of approximating these characteristics to some extent were already available when the need for reducing crosstalk by impedance correction appeared. The first and longest step in this direction was made by O. J. Zobel with his invention of "m-type" sections.[1] The schematic of a typical low-pass [2] filter terminated with these sections is shown in Fig. 6. The terminating networks are enclosed by the

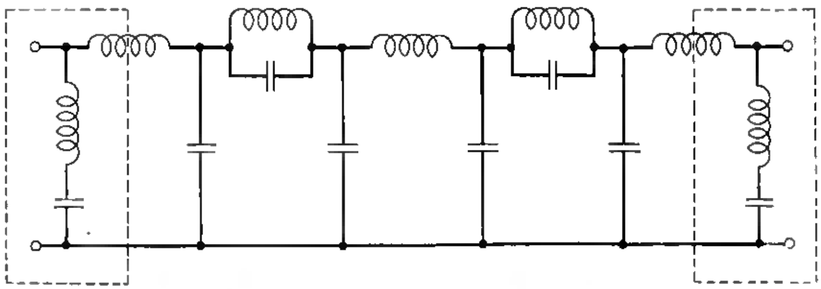


Fig. 6—A typical low-pass filter terminated in "*m*-derived" sections; $m = .512$. The reflection coefficient of this filter is given on Fig. 7.

broken lines. At one end these sections match the normal filter image impedance, as in Fig. 4-*A*. The approximation at the other end to the ideal block type characteristic of Fig. 4-*A* is shown by Fig. 4-*C*. The actual reflection coefficient of the filter of Fig. 6 is given on Fig. 7.

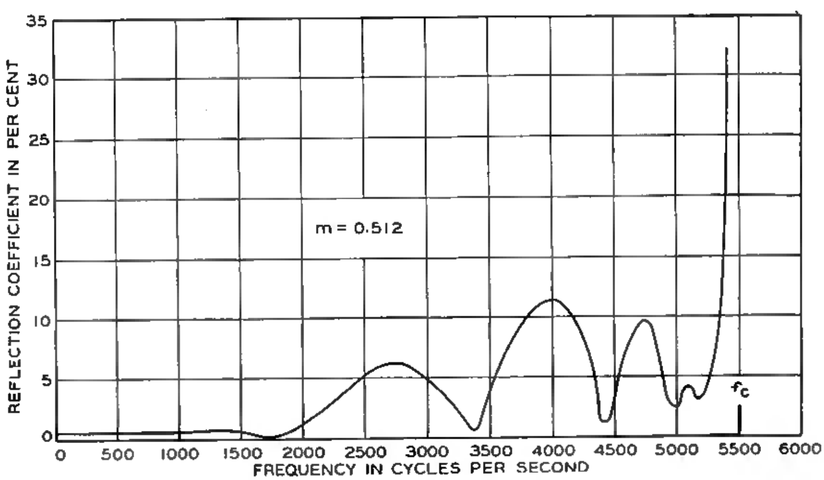


Fig. 7—Reflection coefficient of low-pass filter shown on Fig. 6.

[1] See *Bell System Technical Journal*, Jan. 1923.
[2] The m-type sections are applicable to all types of filters, low-pass, high-pass, or band-pass, and give very similar results in all cases. For example, the curve of Fig. 4-*D* can be considered as being a combination of two curves like that of Fig. 4-*C*, with a slight distortion of the frequency scale. If we allow for this distortion in scales the approximation to the ideal characteristic over a given percentage of the transmitted band is the same for low-pass and band-pass filters.

A modification of m-type sections, leading to the so called x-terminations,[3] is used when filters must be connected in parallel. The modification consists essentially in the elimination of the final shunt branches of the m-derived sections at the paralleling junction. Their places are taken in the transmitting band of either filter by the impedance of the

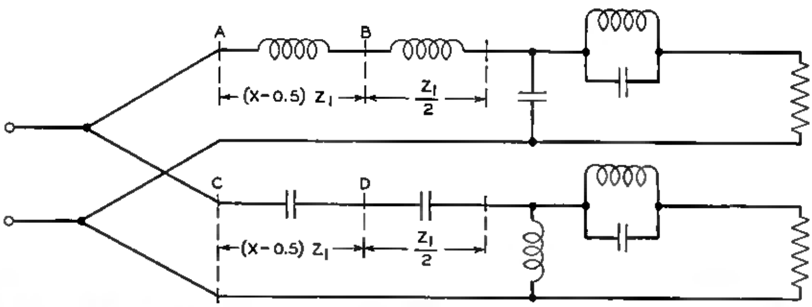


Fig. 8—Schematic of x-terminated filters—showing the way in which the parameter "x" determines the impedance which is added to each filter.

attenuating filter. A simple combination of low-pass and high-pass filters, having x-terminations at their common junction and m-type sections facing their load impedances is shown in Fig. 8. The terminating network for the low-pass filter consists of the impedance AB and that of the high-pass filter while the network for the high-pass filter

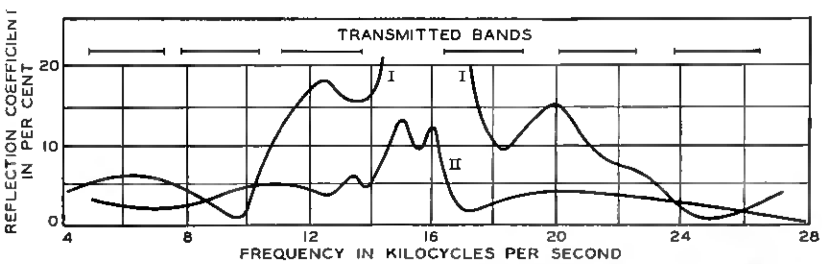


Fig. 9—Reflection coefficient characteristic of parallel low-pass and high-pass filters from the type "C" carrier telephone system.
I—Using x-terminations.
II—After the addition of a simple correcting network to the x-terminated filters.

is the impedance CD and that of the low-pass filter. The reflection coefficient of a typical pair of low-pass and high-pass filters from the Type C carrier telephone system, terminated similarly to the filters of Fig. 8, is shown by Curve I of Fig. 9.

These methods were supplemented by a number of more or less empirical schemes. For example, x-terminations, as Zobel described

[3] See U. S. Patent No. 1557230, issued to O. J. Zobel.

them, could be used only with complementary filters (i.e. low-pass and high-pass, or band-pass and band-elimination). In U. S. Patent No. 1616193 R. H. Mills specifically applies the method to band pass filters. Mills, proceeding from the fact that the adjacent sides of two

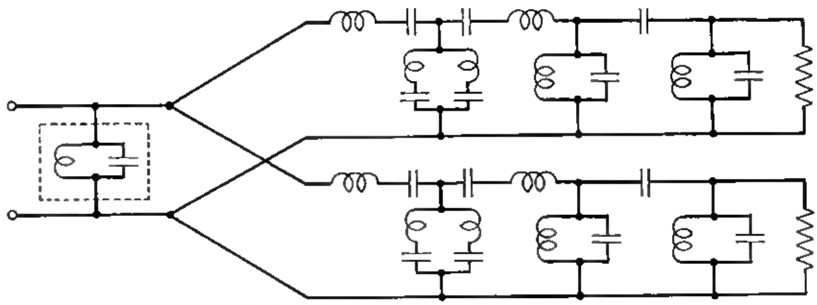


Fig. 10—Schematic of band-pass filters with auxiliary network. The reflection coefficient of these filters is given by curve I of Fig. 11.

band pass filters behave somewhat like complementary filters, while complementary filters are absent on the further sides of the bands, shunted simple networks, approximating the impedances of the missing complementary filters, across the parallel filter system. The filters of the Type *D* carrier telephone system[4] incorporated this device.

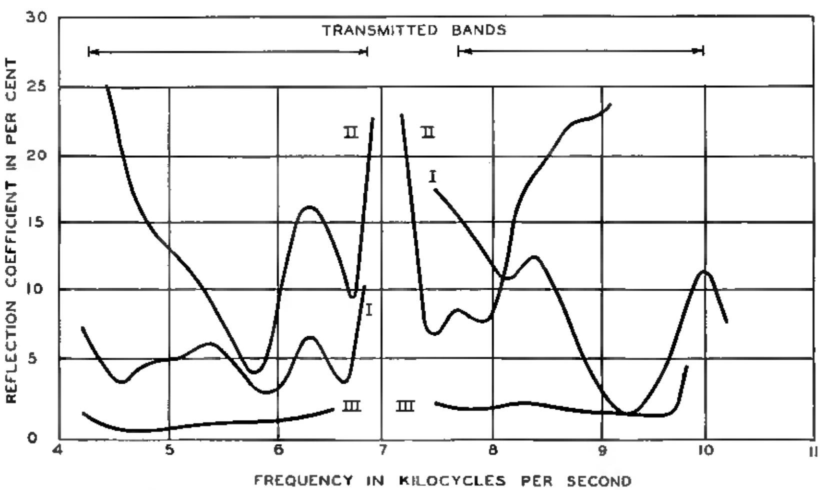


Fig. 11—Reflection coefficient of a set of band-pass filters.
I—For partially corrected x-terminated filters.
II—For uncorrected x-terminated filters.
III—For filters using the termination of Fig. 12-B modified for parallel operation.

[4] The general engineering features of this system are discussed in the *Transactions of the A. I. E. E.*, Vol. 48, No. 1, pp. 117–139.

The filter schematics are shown on Fig. 10, the auxiliary network being enclosed by broken lines. The performance of the filters was further improved by choosing terminating resistances differing somewhat from the nominal or mid-band, value of the filter image impedance. As a result of these two modifications the reflection coefficient characteristic shown by Curve I of Fig. 11 was obtained. Without them, the reflection coefficient would have been that given by Curve II.

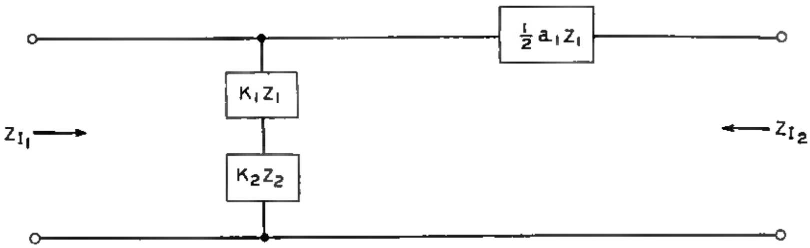
The great improvement of filter impedance characteristics resulting from these devices is evident from a comparison of Figs. 7, 9 and 11 with Fig. 2. Instead of the reflection coefficients of 50 per cent or 60 per cent found in the earliest filters, the technique allows us to obtain reflection coefficients of the order of 10 per cent, within the frequency range of interest, for filters operating alone, of about 15 per cent for pairs of complementary filters in parallel, and of about 20 per cent for systems of parallel band-pass filters. These results were satisfactory for several years. The continued evolution of carrier systems toward higher and higher energy levels, however, and the constant increase in the number of systems in intimate physical association with one another, gradually made such standards inadequate. The reflection coefficient standards demanded by the severe crosstalk requirements of these systems have ranged from 2 per cent to 10 per cent in recent filter designs. It became evident some years ago that if these stringent reflection coefficient requirements were to be met a new analytical technique, more general and more powerful than its predecessors, would be necessary.

A New Technique and the Results of its Application to Impedance Correction

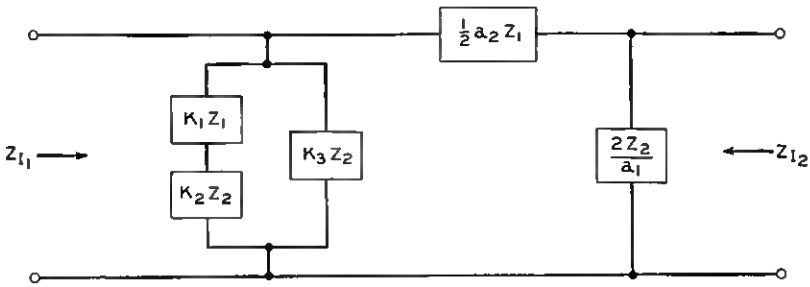
Such a technique has been developed. The method is essentially a generalization of the processes by which Zobel's "x-terminated" filters were derived. It leads to a series of filter sections, the number of which can be extended as far as is necessary to secure a satisfactory approximation to the desired image impedance characteristic.[5] The generalized configurations of several sections are given on Fig. 12. The a's and k's of this figure are design parameters, Z_{1k} and Z_{2k} refer to the filter with which the section is to be used. By choosing Z_{1k} and Z_{2k} appropriately the terminations can be adapted to any type of filter structure, whether low-pass, high-pass or band-pass.

The simplest of these sections can be shown to be equivalent to an "m-type" structure, and will naturally give the same results. A

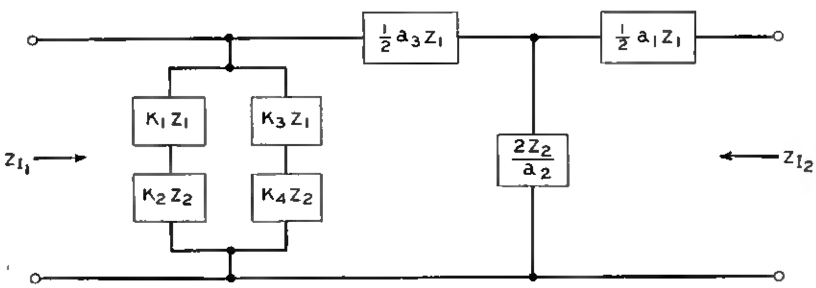
[5] For a detailed discussion of the theory underlying this technique see, "A Method of Impedance Correction," appearing simultaneously in this journal.



A



B



C

Fig. 12—Generalized schematics of terminating sections.

A—An "*m*-derived" or "single-branch" network.
B—A "2-branch" network.
C—A "3-branch" network.

typical image impedance [6] characteristic of the next more complicated structure (Fig. 12-*B*) is shown on Fig. 13. The vertical scale of this figure has been made considerably larger than that of Fig. 4 in order

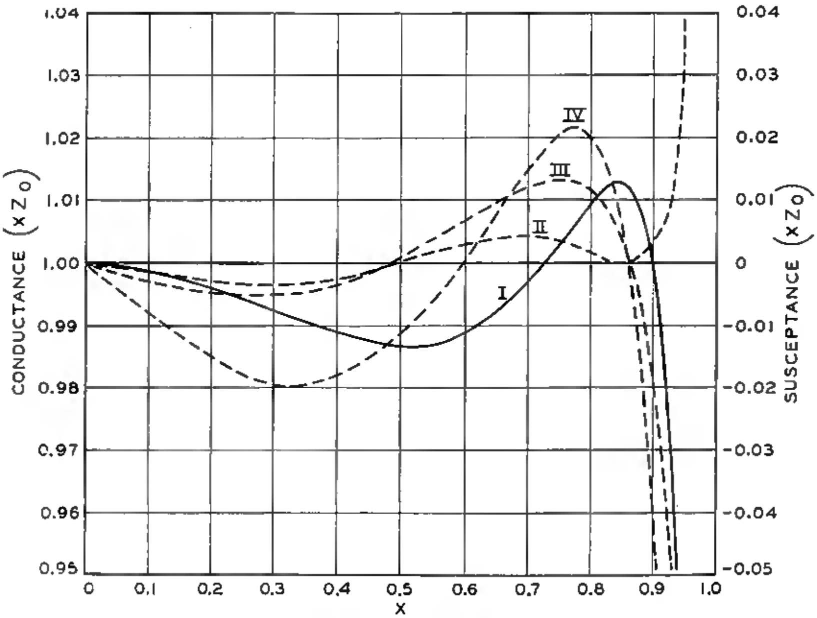


Fig. 13—Typical image impedance characteristic of a 2-branch termination (Fig. 12-*B*).
I—Real component.
II, III and IV—Various possible imaginary components.

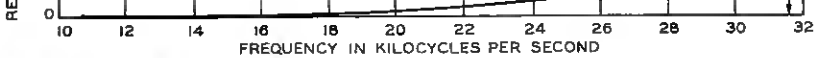


Fig. 14—Schematic and reflection coefficient of a low-pass filter with 2-branch terminations.

[6] The fact that the so-called "image impedances" of Figs. 12 and 16 contain slight imaginary components, in defiance of the fact that the image impedance of a reactive network is never a complex quantity, is traceable to the method used in analyzing the structures, which will be more fully understood from the discussion in the accompanying paper. The curves actually represent the impedance of the section when they are terminated in the filter image impedance. The reason for showing several curves for the imaginary component will be brought out later.

to bring out the departure of the image impedance from its ideal value more clearly. A low pass filter to which the termination has been applied is shown on Fig. 14. The terminating sections are enclosed by the broken lines. The resulting reflection coefficient is given on the same figure. It will be observed that the reflection coefficient over 93 per cent of the transmitting band is less than 2.5 per cent, or only about one-fourth that of the analogous m-type terminated filter shown in Fig. 6. The application of the structure to band-pass filters is illustrated in Fig. 15, which represents a portion of the redesigned type

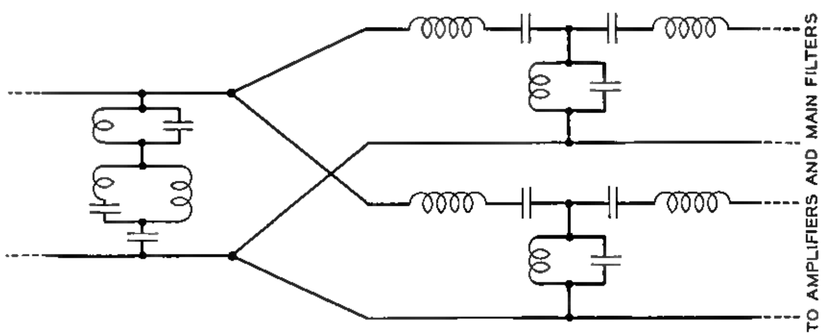


Fig. 15—Band-pass filters with 2-branch terminations modified for parallel operation. The reflection coefficient of these filters is given by Curve III, Fig. 11.

"*D*" system. The configuration shown in Fig. 12-*B* has been modified in these networks to adapt the filters for parallel operation. There are no separate terminations at the receiving ends of the filters since with these very simple structures reflection at the receiving ends could be taken into account by a slight adjustment of the terminating network facing the line. The improvement produced by the new networks in the reflection coefficient characteristics of the system is evident from a comparison of Curves III and I of Fig. 11.

The terminating section shown in Fig. 12-*B* which is one step beyond the "m-type" section in complexity, is adequate in most situations. When a severe reflection coefficient requirement must be met almost up to the filter cutoff, however, it is necessary to resort to the more complicated configuration of Fig. 12-*C*. The image impedance characteristic of this section, when its parameters are adjusted for an operating range extending over 97.5 per cent of the theoretical transmitting band, is shown on Fig. 16. The reflection coefficient actually obtained when sections of this type, but with somewhat different values of the design parameters, were applied to a high-pass filter is shown by Fig. 17. The filter configuration is given on the same figure. In this figure the susceptance controlling network of the high-pass filter at the paralleled end is composed of the coil and condenser in series across the input terminals and the susceptance of the low-pass filter; at the other end

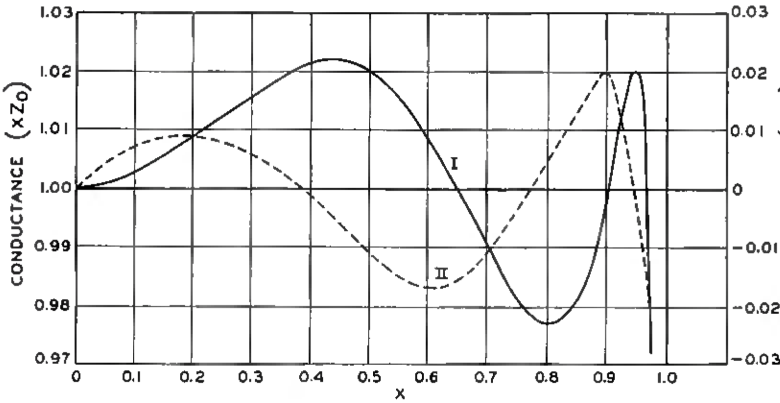


Fig. 16—Typical image impedance characteristic of a 3-branch termination (Fig. 12-*C*).
I—Real Component.
II—Imaginary component.

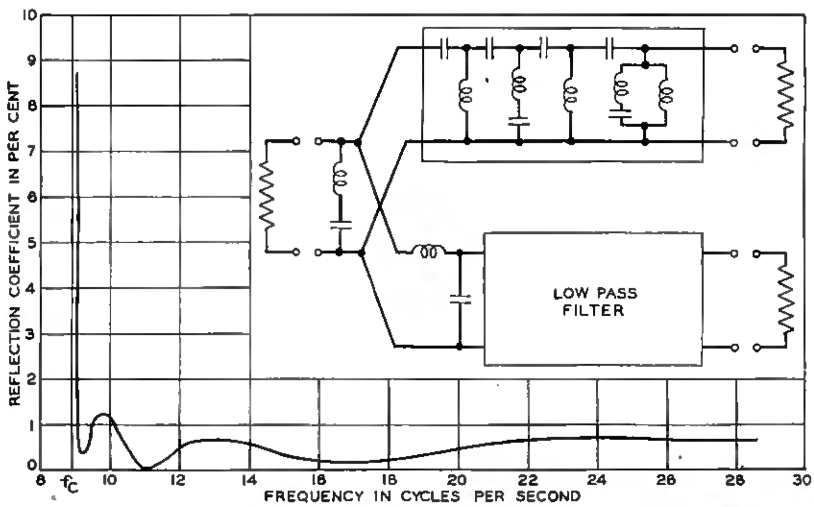


Fig. 17—Schematic and reflection coefficient of a high-pass filter with 3-branch terminations.

of the high-pass filter the three-element two-terminal network controls the susceptance. The conductance controlling sections at either end are composed of the first series condenser and the first shunt coil and a portion of the second series condenser. The maximum reflection coefficient over about 95 per cent of the nominal transmitting band is slightly greater than 1 per cent. We can summarize these quantitative results in the rough statement that each of the three stages in the progress from the most primitive filter section to the relatively complicated network of Fig. 12-*C* appears to reduce the reflection coefficient obtainable over a given frequency range by a factor of about three or four.

Impedance Correction for Filters Operating in Parallel

The modifications which must be made in these sections in order to adapt them for use with filters which must operate in parallel are similar to those which were made in adapting "m-type" sections to this service. The final branch of each termination is omitted, its place being taken within the transmitting band of the filter to which it belongs, by the impedance of the parallel, attenuating, filters of the system. The parallel filters cannot however be relied upon to simulate the missing branch, even in this frequency range, with great accuracy. If we wish to preserve the high standards achieved by the terminations in other circumstances, therefore, it is, in general, necessary to introduce an auxiliary network in shunt with the circuit as a whole to improve the approximations to the missing branches. When this is done the reflection coefficient of the complete system is substantially identical in any transmitting range with that which would be obtained from the corresponding filter operating alone.

The thorough exploitation of the possibilities of these auxiliary networks leads to a marked improvement in the performance even of the well-known x-terminations. The reflection coefficient characteristic of a typical pair of high- and low-pass filters has already been shown by Curve I of Fig. 9. The high value of the reflection coefficient of these filters is largely due to the fact that neither filter in its attenuating range supplies quite enough admittance to take the place of the missing shunt branch of the other filter. The addition of a simple tuned circuit resonating between the transmitting bands to compensate for this deficiency in admittance reduces the reflection coefficient to the level shown by Curve II. The results for x-terminated band-pass filters are even more striking. Fig. 18 gives the susceptance at the line terminals of a set of three filters for several different conditions. The susceptance should ideally be zero. Curve I gives its value when no auxiliary network is added, Curve II, the level to which it is reduced by the auxiliary network suggested by Mills, and Curve III the characteristic which can be obtained with the help of a more elaborate auxiliary network. These curves can be given quantitative significance if we notice that the deviation represented by Curve I at the point marked by the arrow would lead to a reflection coefficient of about 50 per cent even if the system were otherwise ideally terminated. Curves II and III, under the same conditions, represent reflection coefficients of about 29 per cent and 3 per cent respectively.

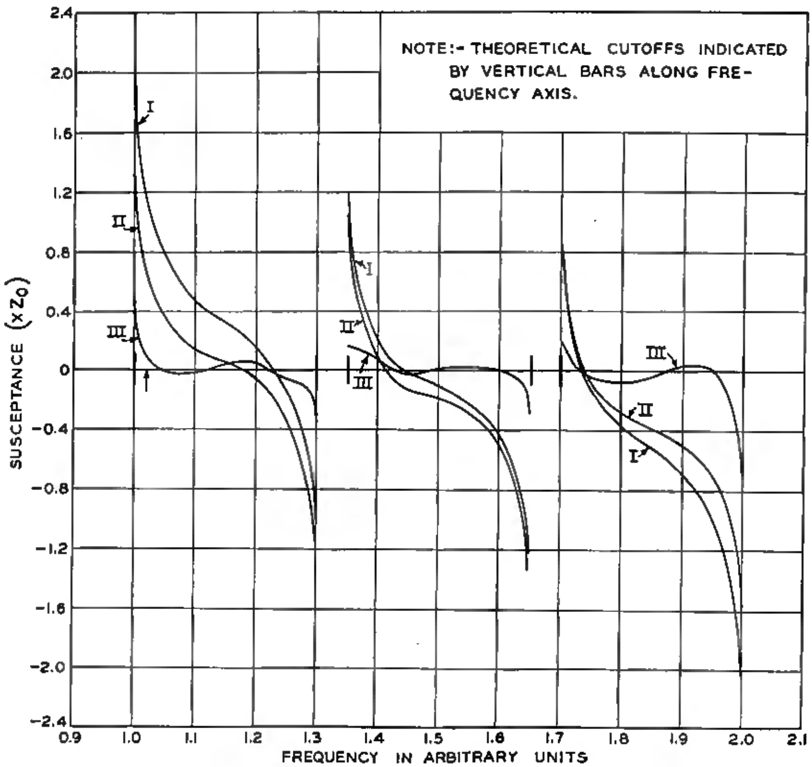


Fig. 18—Susceptance correction of x-terminated band-pass filters.
I—Uncorrected Susceptance.
II—Susceptance after the addition of a simple auxiliary network.
III—Susceptance after the addition of a more elaborate auxiliary network.

Improvements in Filters for Use with Modulator and Demodulator

We have hitherto restricted our attention to the impedance characteristics of filters within their transmitting bands since it is only in this range that impedance irregularities in the circuit can produce crosstalk. When a filter operates in conjunction with a modulating device, however, a high modulator efficiency with low distortion demands that the impedance of the filter to the untransmitted side band be low (or high) and nearly constant. All of the correcting networks we have thus far described produce sharp changes in reactance of the attenuating region and are therefore unsuitable for such circuits. In spite of their poor characteristics within the transmitting band, therefore, it has hitherto been necessary to use mid-shunt image impedance terminations of the primitive "constant-k" type at junctions between filters and modulators. Curve I of Fig. 19 for example, shows the

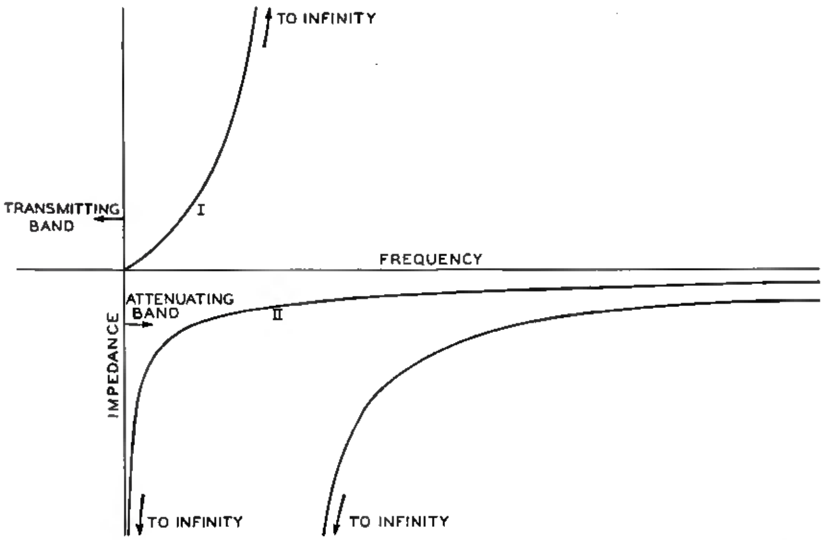


Fig. 19—"Constant-k" and "m-derived" type image impedances in the attenuating range.
I—"m-derived" type image impedance.
II—"Constant-k" type image impedance.

impedance characteristic of an "m-type" section beyond the cutoff in comparison with Curve II, representing the impedance of the "constant-k" type section in this range. The network configurations shown in Fig. 12 however represent only one of two possible classes of sections which can be developed as a result of the general analysis given in the accompanying paper. The other class is radically different in configuration. Networks of this second class may be advantageously substituted for the "constant-k" sections formerly used with modulators. The network characteristics in their attenuating regions approximate those of the "constant-k" sections and while they are not quite as good in the transmitting region as the characteristics furnished by the networks of Fig. 12 they are much better than the characteristics of the "constant k" sections of Figs. 4-*A* and 4-*B*.

Attenuation of Impedance Correcting Sections Reduces Net Cost of Impedance Correction

The economic aspects of filter design demand some sort of an evaluation of the cost of improving filter impedances. While the terminat-

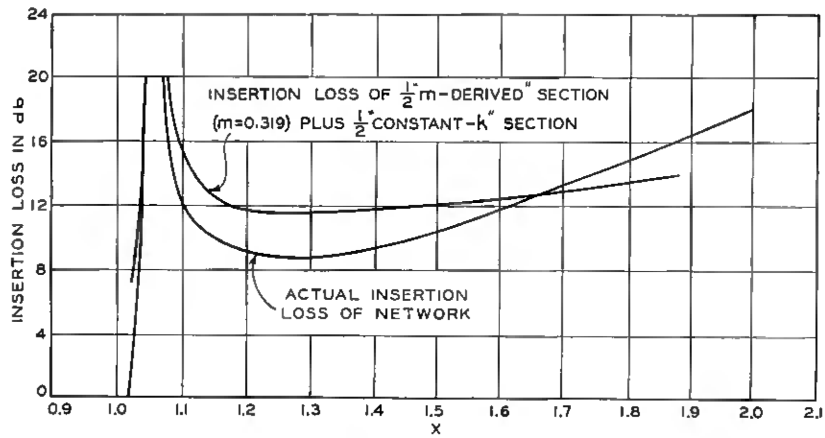


Fig. 20—Insertion loss of a 2-branch termination.

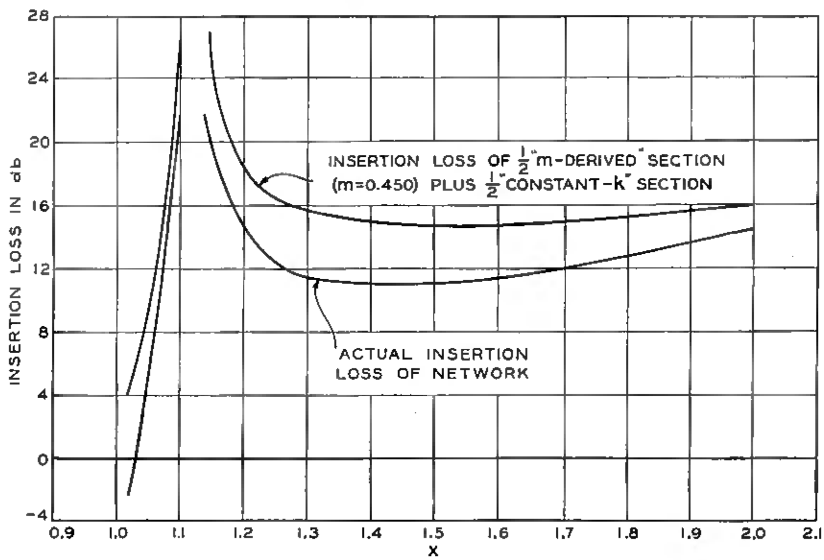


Fig. 21—Insertion loss of a 2-branch termination.

ing sections shown in Fig. 12 are rather complicated, their cost is discounted considerably by the fact that they contribute appreciably to the attenuation of the structure as a whole to undesired frequencies.

Moreover, their attenuation characteristics can be varied within fairly wide limits without appreciably affecting the impedance characteristics we obtain. If we make allowance beforehand for the attenuation of the terminations, therefore, the number of sections making up the main body of the filter can be correspondingly reduced. These relations are illustrated by Figs. 20, 21 and 22 which are drawn for terminations having the configuration of Fig. 12-*B*. The network attenuation is compared in each case with the attenuation of the most nearly equiv-

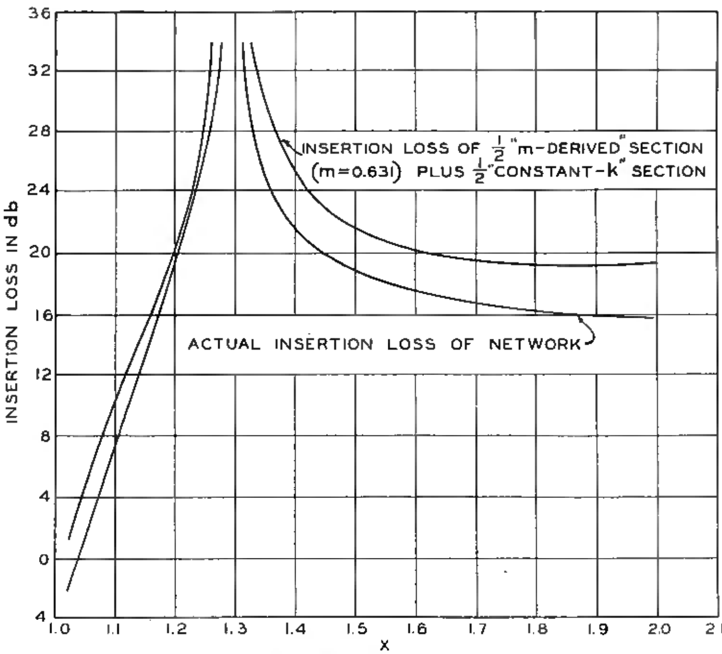


Fig. 22—Insertion loss of a 2-branch termination.

alent filter structure. The corresponding modifications in the impedance characteristic of the network are shown by Curves II, III and IV of Fig. 13. The real component (Curve I) of the impedance is the same in all cases, since the adjustment of the attenuation characteristic was produced entirely by manipulating the final series branch of the network, which has no effect on this component. When low- and high-pass filters are involved the terminating networks contain one more element than the suggested filter equivalent. This much must be conceded to the cost of impedance correction. It will be observed, however, that the remaining elements contribute almost as much attenuation as they would in standard filter sections. Indeed at frequencies remote from the cutoff the attenuation of the network considerably exceeds that of the filter equivalent. The attenuation produced by the auxiliary (susceptance) networks used in conjunction with parallel filters is not so easily evaluated in terms of a standard filter equivalent. Since these networks produce peaks of attenuation just beyond the filter cutoff, thus enhancing the selectivity of the systems, they are however, in some respects particularly valuable. We can summarize the economic aspects of impedance correction in the statements that a severe impedance requirement will increase the number of elements (coils and condensers) required for an average filter used in carrier circuits by about 15 per cent or 20 per cent, and that the corresponding increase in the cost of the filter as a whole will be about 10 per cent or 15 per cent.

Practical Limitation to Impedance Correction

The fundamental limitation on the correction of wave filter impedances is practical rather than analytical. In other words it depends upon the accuracy with which it is possible to manufacture filters. All the curves so far exhibited have been based on the assumption that the filter elements, coils, condensers, and resistances, have the

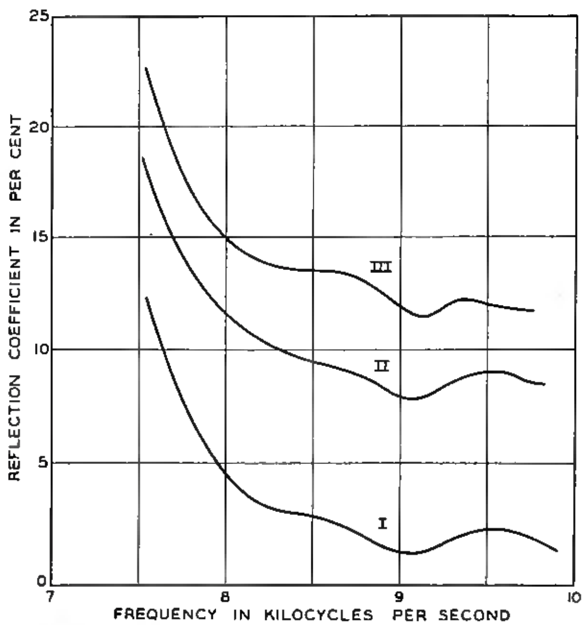


Fig. 23—Effect of element variations on a filter reflection coefficient.
I—Reflection coefficient when all elements have their exact design values.
II—Envelope of reflection coefficients of the best 99 per cent of the filters.
III—Envelope of worst possible reflection coefficients.

exact values ascribed to them by the design formulæ. The limitations of manufacture, however, demand that elements be permitted to deviate from their mean values by as much as 1 per cent or 2 per cent. These element deviations in general degrade the performance of filters by increasing their reflection coefficients. The increase in reflection coefficient is largely independent of the initial reflection coefficient, that is, it is about the same for a filter whose normal reflection coefficient is very small as for one of inferior theoretical design having a large normal reflection coefficient. Curve I of Fig. 23 shows the reflection coefficient of a particular filter when all of the elements have their design values; Curve III, the envelop of maximum reflection coefficients for this filter

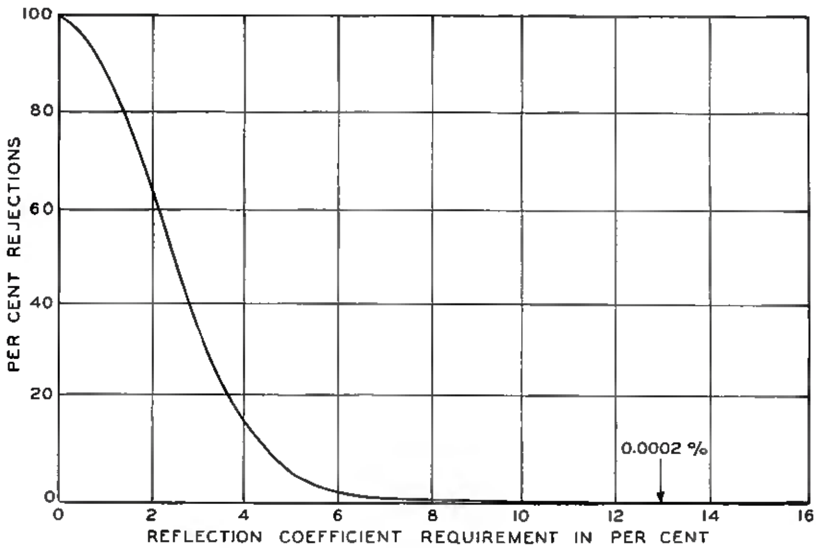


Fig. 24—Distribution of the reflection coefficients of a given filter at one frequency in terms of the percentage rejected in meeting any requirement.

when the coils are permitted to vary ± 1.5 per cent and the condensers ± 0.8 per cent from their design values; Curve II the maximum reflection coefficient for the best 99 per cent of the filters manufactured to these limits. The last curve is based on a priori probability computation which assumes that the distribution curves of the element deviations follow the normal law. A distribution curve of reflection coefficients for another filter at a particular frequency plotted in terms of the per cent of filters that would be rejected in meeting any requirements is given in Fig. 24. Studies of this sort are too extensive to be included in this paper but they have shown that the advances in the technique of impedance correction here recorded are well ahead of the practical limitations of the problem.

A Method of Impedance Correction

By H. W. BODE

This paper gives a theoretical treatment of some recently developed wave filter terminating sections whose application is discussed in the accompanying paper on "Impedance Correction of Wave Filters." The sections consist primarily of non-recurrent ladder networks which operate, over the transmission bands of the associated filters, as transformers whose ratio varies with frequency. The transformation ratio of the network is specified, as a function of frequency, by a power series containing a limited number of terms and the design procedure therefore depends upon the construction of power series approximations to the ratio between the resistance of the filter proper and the desired resistance. A separate network is added to secure control of the reactance component. An increased number of terms in the power series, and therefore an improved approximation to the desired transformation ratio, can be obtained by increasing the number of branches in the network. The method thus leads to a series of sections of progressively increasing complexity and with progressively improving impedance characteristics. By an inversion of the analysis a second series of sections can also be obtained. The paper is chiefly devoted to a discussion of these two series of filter sections, but other possible applications of the method are also described briefly.

THE analysis of transmission circuits with which telephone engineers are familiar is an outgrowth of the general physical theory of the propagation of wave disturbances in continuous media. Problems analogous to the analysis of a smooth transmission line are found, for example, in optical and acoustical theory and in the theory of the vibrations of a taut string. The situations of most importance from the standpoint of general physics are those in which the continuous medium extends indefinitely in at least one direction. Since, moreover, this is also the simplest case, it has been customary to base our transmission analysis upon the analogous concept of an infinite line with distributed constants. The analysis of such a structure, since it depends upon only two quantities, the characteristic impedance and the propagation constant, is of course very simple.

An actual telephone transmission circuit, however, is by no means an infinite structure containing distributed constants. Many lines, for example, are loaded. Whether loaded or unloaded, they do not extend indefinitely, but are interrupted by terminal apparatus and intermediate repeaters. Each of these, moreover, contains a miscellany of apparatus, such as modulators, transformers, amplifiers, filters, equalizers, by-pass circuits, and the like, having little physical resemblance to a continuous medium.

This physical contrast between an ideal continuous medium and an actual physical telephone circuit does not necessarily mean that the application of the wave theory to circuit analysis is a difficult matter. To a first approximation we can determine the response of a circuit merely by adding together the propagation constants of its various constituents. Unfortunately, however, the diverse components of a typical circuit usually have characteristic impedances which are widely different functions of frequency. Thus, for example the impedances of the amplifiers and modulators in most telephone systems are nearly constant pure resistances. Non-loaded lines approach such a characteristic at high frequencies but at low frequencies their impedance is usually large and may have a considerable reactive component. Loaded lines depart from a constant resistance at high frequencies as well. An even more complicated characteristic, consisting of a varying resistance in the transmitted band, changing abruptly to a pure reactance as we pass the cutoff, is exhibited by a wave filter. In addition to the normal propagation constants of the circuit, therefore, we must take account of reflection effects at all of the junctions between these various types of characteristic impedance. In a long circuit containing impedance irregularities at many junctions, moreover, we must give consideration to an enormous variety of waves which suffer multiple reflections from a number of junctions. This complicated system of factors may make life burdensome to the man who must evaluate them, but since they are seldom large enough to grossly affect the transmission characteristic of a circuit, they usually play otherwise a secondary role in practical transmission analyses. They do, however, blur the original clarity of the wave picture and from the standpoint of theoretical simplicity at least, therefore, they should be eliminated. For this purpose we should have at our disposal a network whose impedances at its two ends could be assigned arbitrarily to match the impedances actually present at any junction.

The networks which form the subject of this paper were developed to eliminate reflection effects which, in addition to being a nuisance from the theoretical standpoint, were attended by serious practical consequences as well. The engineering problem involved is described in the paper on "Impedance Correction of Wave Filters" by E. B. Payne appearing simultaneously in this *Journal*. Briefly, it appears from the discussion in that paper that impedance mismatches at the junctions between terminal or repeater equipment of carrier systems and the line give rise to reflected waves which may produce cross-talk in neighboring systems. This cross-talk can be reduced as much as we like by means of line transpositions but the required transposition scheme is so extremely complicated and expensive that the reduction in the amplitude of the reflected waves by improvement of the reflection coefficient at these junctions is of considerable economic importance. The impedances of the terminals and repeaters at the junctions at which reflections occur are chiefly determined by their filters, which are the apparatus immediately facing the line. A detailed study of the relationship between the actual input impedance of a filter and mismatches of characteristic impedance which may occur at further junction points in the circuit shows that by far the simplest method of obtaining a low reflection coefficient at the line terminals is to produce a match of characteristic impedances at all junction points of the filter system. Fortunately speech currents beyond the transmitted band of the filters carry so little energy that the reflection coefficient of the structure in these ranges is of no importance. The technical problem therefore reduces to the construction of a new type of filter section for use at terminations, the new filter section having an image impedance within the transmitted band which at one end matches that of the standard sections forming the main body of the structure and at the other approximates a constant resistance, matching the terminating impedances. Of course the new filter sections must also be so chosen that they will not impair the transmission properties of the system.

This immediate problem has been solved. It still leaves unsettled, however, the question as to whether we can devise a type of network capable of correcting for reflection effects not only at these particular junctions but also at any other impedance irregularity in the circuit. Such a structure would transform one arbitrary impedance characteristic into another preassigned characteristic without decreasing the transmission efficiency of the circuit, much as the familiar attenuation equalizer changes the attenuation characteristic of a circuit by a preassigned amount without changing its impedance and without greatly affecting its phase characteristic. The mathematical analysis underlying the sections which have been developed for filter impedance correction is easily extended to a much broader class of terminating impedances. Judged from a purely formal standpoint, therefore, the networks appear to be a long step forward in the development of such a general impedance equalizing device. Unfortunately, it seems certain from other considerations that much of the promise thus inherent in the formal mathematical analysis may not be realized in practical applications, but since the network has been thoroughly studied only in its application to filters, its precise limitations are still uncertain. In the discussion which follows the general method of impedance correction is first sketched briefly, and is succeeded by a detailed treatment of its application to filters. Some of the probable limitations of the method in other applications are suggested near the end of this paper.

The analysis used in impedance correction can also be applied to the construction of networks having transmission properties somewhat like those of the familiar wave filter. In contrast to the usual filter theory, developed, after the analogy of wave propagation in continuous media, from the conception of an infinite recurrent structure, however, it leads to networks which are not recurrent and are not divisible into separate sections with matched image impedances. In its present state of development the analysis is unquestionably much less powerful than the established theory. Since it may be of interest as an example showing at least the possibility of an alternative approach to filter design, however, it is discussed briefly at the conclusion of the paper.

General Impedance Correcting Process

If no transmission requirements were imposed upon electrical structures, a wide variety of networks might be used for impedance correction. For example, we might make up deficiencies of impedance or admittance by a simple two-terminal network in series or in shunt with the circuit. In almost all circuits, however, we are interested in securing minimum transmission loss, that is to say, maximum energy in the receiving impedance, throughout the frequency bands containing the transmitted signals. The energy which goes into a system terminated by a correcting network depends only upon generator and the corrected impedance, both of which are specified by the conditions of the problem. We can increase the energy delivered to the receiving device, therefore, only by reducing the amount absorbed in the correcting network. Obviously the best possible condition is found when the correcting network is composed of pure reactances. Unless either the resistance or the conductance of the circuit happens to be ideal, however, impedance correction cannot be obtained by a simple two-terminal reactive network. For this reason, the impedance correcting structures which have been developed are four-terminal networks of pure reactances. Control of the resistance or conductance component is gained, not by the direct addition of resistance, but rather through the use of the network as a sort of variable transformer, whose impedance ratio changes as we go over the frequency range. In such a circuit the insertion loss of the network is determined entirely by the ratio of the energy drawn from the generator by the original and the corrected impedance. Ideal dissipationless network elements are, of course, not available in practice. Except for the possible influence of this factor, impedance correction, since it normally means an improvement in the match between generator and load impedances, should result in a slight increase of transmission efficiency.

Reciprocal Impedance Relations at Terminals of a Reactive Network

Our restriction to networks of pure reactances allows us to make use of a principle by means of which the impedances measured at the two ends of the network under certain terminal conditions can be reciprocally related to one another. The theorem will be given here since it is of frequent application in further discussion. Referring to Fig. 1, let us assume that the impedance measured at terminals *cd*, with an impedance $\overline{Z}_1$ connected to terminals *ab*, is equal to $\overline{Z}_2$, as is shown in the diagram. The theorem is concerned with the impedance Z looking into terminals *ab* when Z_2, the conjugate of $\overline{Z}_2$, is connected across *cd*. Let us suppose that the generator e in Z_2 produces a current i in $\overline{Z}_1$. Then, by the usual principle of reciprocity, the generator e when inserted in $\overline{Z}_1$ will produce the current i in Z_2. In the first case the power entering the network is obviously $\frac{e^2}{4R_2}$ and the power flowing from it into $\overline{Z}_1$ is i^2R_1. In the second case these powers are $\frac{e^2R}{(R_1+R)^2+(X_1+X)^2}$ and i^2R_2. Since the network is non-dissipative the power entering the network equals the power leaving it in both cases.

$$\frac{e^2}{4R_2} = i^2R_1$$

$$\frac{e^2R}{(R_1+R)^2+(X_1+X)^2} = i^2R_2$$

Upon dividing the two equations and simplifying we find:

$$(R_1 - R)^2 + (X_1 + X)^2 = 0$$

which can be true only if:

$$R = R_1 \text{ and } X = -X_1$$

In other words, Z is the conjugate of $\overline{Z}_1$. We can state this result in the following words:

A network composed of pure reactances will have a given impedance, Z_1 at one pair of terminals when an impedance Z_2 is connected to a second set of terminals if, when the conjugate of Z_1 is connected to the first pair of terminals, the impedance measured at the second pair of terminals is the conjugate of Z_2.

This theorem can be applied immediately to the problem of filter impedance correction discussed in the introduction. The networks required for this problem were defined there as sections which within the transmitting band would have image impedances matching the line at one end and matching the image impedance of the main body of the filter at the other. If we represent the filter proper and the line by $\overline{Z}_1$ and Z_2 in Fig. 1, these image impedance requirements reduce to the

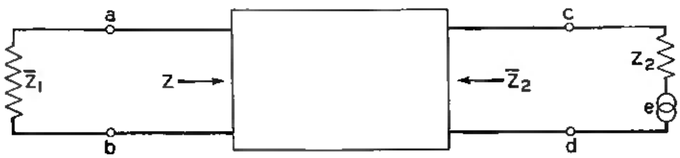


Fig. 1—Diagram to illustrate the reciprocal properties of impedance correcting networks.

statement that the network must be so chosen that an impedance match exists both at *ab* and at *cd*. $\overline{Z}_1$ and Z_2 for this particular circuit are however, pure resistances, and therefore equal to their conjugates, within the required frequency range. The theorem shows that an impedance match will be obtained at *cd* provided an impedance match exists at *ab*, and vice versa.[1] If we please, therefore, we can consider that our problem is that of obtaining a network which, when terminated by a filter, has an actual impedance equal to a constant resistance. On the other hand we can start with the resistance and attempt to build up a network whose impedance matches that of the filter. Both the first or "direct" and the second or "reverse" methods of constructing terminating networks for filters are considered in the next section. With either procedure the resulting networks have both required image impedances and can be used, when properly connected, either at the line or the receiving end of the filter. The "correction" of one impedance to match another and the construction of a network having given image impedance characteristics are therefore interchangeable conceptions.

Separate Correction of Real and Imaginary Components of Impedance or Admittance

The image impedance method of defining the properties of the terminating network is a convenient one when we are concerned with the operation of the structure in the transmission system as a whole. The methods used in designing the network can, however, be described

[1] See also Feldtkeller's paper, "Über einige Endnetzwerke von Kettenleitern" in the *Elektrische Nachrichten-Technik*, June 1927, for a very similar use of this property of reactive networks.

more simply if we reject the image impedance statement of the problem in favor of its alternative. For the time being, therefore, we will assume that we are attempting to design a reactive network having a preassigned input impedance when terminated by a given load impedance.

In accordance with a principle originally stated by O. J. Zobel,[2] this problem of impedance correction can be simplified if we consider separately the resistance and reactance of the corrected circuit. To be more explicit, since a reactance in series with the circuit will change its reactance without changing its resistance, it is simplest to consider first the construction of a network which will produce the required resistance characteristic. Of course the reactance characteristic furnished by such a structure will not in general be ideal, but we may be able to correct it to the proper value by the later addition of a series reactive network. Quite obviously, it is equally easy to base the analysis upon admittances and construct first a network which will give the required conductance characteristic and make up any faults in its susceptance characteristic by a final shunting branch.

This division of the network into two separate structures is, of course, not a necessary one and in view of the extremely limited range of reactance or susceptance characteristics which can be compensated for by a final, physically realizable, two-terminal reactive network may seem scarcely desirable. An alternative procedure in which this division is not attempted is mentioned in the concluding section. The reason for assuming separate correction of the real and imaginary components of impedance and admittance in the present discussion is simply one of convenience. The difficulties which might be anticipated in the design of the final reactive compensator do not appear in filter impedance correcting problems, at least. On the other hand, the division has the advantage that it makes each step simple and allows us to meet fairly severe impedance requirements with a small number of variables. As we shall see later the method has the further advantage in its application to filters that it lends itself readily to the modifications necessary when a number of filters must operate together.

The Resistance or Conductance Controlling Network

Since the characteristics of two-terminal reactance networks are well understood, the construction of the final reactive branch demands no

[2] See, for example, U. S. Patents No. 1,557,229 and 1,557,230 where he applies it to "x-terminated" filters. The method of this paper is in some respects merely a generalization of that analysis. The relation of "m-derived" sections and "x-terminations" to the filter terminations developed in this paper is indicated in the following section. In this connection, the previous work of R. S. Hoyt on loaded lines should also be mentioned. See this *Journal*, Vol. 3, p. 414, 1924.

extensive discussion. The problem of designing a four-terminal reactive network which will transform one arbitrary resistance or conductance characteristic into another arbitrary characteristic must, however, be treated with more respect. The configuration which has been adopted for this purpose is shown in Fig. 2. The quantities of

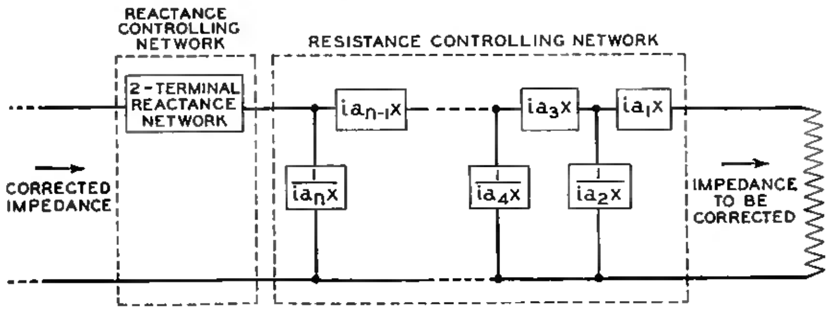


Fig. 2—Generalized schematic of impedance correcting network.

the general form ia_jx are analytic representations of the impedances of the series branches and admittances of the shunt branches. The a's are constants whose choice determines the particular resistance or conductance controlling properties of the structure, and x is a function of frequency. Since the series impedances and shunt admittances are all proportional to x all of the series branches will have a given physical configuration and all of the shunt branches will have the inverse configuration. For example if the series branches are inductances the shunt branches will be capacities, while x, of course, will be proportional to frequency. By using other series arm configurations we can obtain a considerable variety of networks. Each such network, it will be noticed, is similar to a "constant-k" filter in physical configuration. The appropriate network in any particular situation is that one which resembles a constant-k filter transmitting the frequency range of interest.

The property of this network configuration which makes it particularly suitable as a resistance or conductance controlling device is the fact that in most instances the modification it produces in the resistance or conductance of the load can be expressed as a single polynomial. To be more explicit, when the load impedance is of a certain mathematical type, which includes the impedances in which we are most interested, the resistance or conductance of the corrected structure is given by a formula of the following sort.

$$R(\text{or } G) = \frac{F(x)}{A_0 + A_1x + A_2x^2 + \cdots + A_nx^n}$$

in which the A's are constants involving the arbitrary quantities a_1, a_2, etc. which specify the network elements.

The quantity $F(x)$ is usually either the resistance or conductance component of the load impedance, and in any case is a quantity entirely determined by that impedance. In order to secure the proper resistance or conductance from the corrected structure, therefore, it is merely necessary to choose such values of the constants $A_0 \cdots A_n$ that the polynomial satisfies the equation

$$A_0 + A_1 x + \cdots A_n x^n = \frac{F(x)}{R(\text{or } G)}$$

with sufficient accuracy when R is given the desired value of the corrected resistance or conductance. The problem of approximating a given curve by a polynomial of given degree is a well known one in mathematics and such general methods as expansions in power series or Legendrian harmonics exist for its solution. We can, therefore, consider that the choice of these constants presents no particular difficulty. Even without the help of these general methods, however, the problem is so simple that suitable approximations can be obtained by cut-and-try methods.

These polynomial coefficients $A_0 \cdots A_n$ are merely intermediate parameters which specify the values of the elements in the network implicitly but do not give them directly. In order to determine the relation between these coefficients and the actual element values it is necessary to make a direct computation of the impedance of the network in terms of the a's and sort out the various powers of x in the resulting expression. Each of the quantities $A_1 \cdots A_n$ is thereby expressed as a function of the a's. Our next step must then be to determine values of the network elements by solving the set of simultaneous equations relating them to the numerical values of the polynomial coefficients. In accordance with the procedure we have adopted, the design is completed by the computation of the reactance or the susceptance of the network, and its adjustment to the desired value by the addition of a suitable final branch. The discussion of the application of the method to filter impedances given in the next section will illustrate the process in detail.

Proof of Properties of Ladder Type Resistance Correctors

As we observed in a previous paragraph the ratio of the load resistance or conductance to the corrected resistance or conductance can be expressed in this simple fashion as a polynomial in x only when the load impedance belongs to a certain mathematical class. Appropriate load impedances are those whose real components can be written as the square roots of rational functions[3] of x and whose imaginary components are rational functions of x. We can make this conclusion plausible by direct inspection. It is obvious that the general nature of the mathematical expression for the impedance of the network cannot change radically as we add successive branches. When we add a series branch, however, the reactance is increased by $a_j x$, while the resistance is not altered. The functional form of the impedance then will be unchanged if the reactance was originally an algebraic function of x. But, since we must add shunt as well as series arms to the network the functional forms must be symmetrical whether taken on an impedance or admittance basis. By analogy, therefore, the susceptance also must be a rational algebraic function. The susceptance B is expressed in terms of R and X, the resistance and reactance, by $B = X/(R^2 + X^2)$, but X (and therefore X^2) has already been fixed as a rational algebraic function and R^2 must have a similar form if the whole susceptance expression is to be such a function. This conclusion, since it applies equally at any part of the network, must, of course, be valid for the load impedance also.

This argument is sufficient to indicate what sort of a load impedance *might* have the property for which we are looking—that of allowing the change in resistance or conductance produced by the insertion of the ladder network to be expressible as a simple polynomial. In order to show definitely that this type of load impedance *will* have that property it is simplest to begin by finding out whether the relation holds when the network consists of a single branch. In accordance with the previous discussion, the load impedance will be taken as

$$\sqrt{\frac{F_1(x)}{F_2(x)}} + i\frac{G_1(x)}{G_2(x)},$$

where $F_1(x)$, $F_2(x)$, $G_1(x)$, and $G_2(x)$ are polynomials in x. Upon multiplying and dividing the resistance expression by $\sqrt{F_2(x)}C(x)$, where $C(x)$ is a new polynomial so chosen that when the product $F_2(x)C(x)$ is divided by $G^2(x)$ the quotient is a polynomial, the load impedance is transformed into

$$\frac{\sqrt{F_1(x)F_2(x)C^2(x)}}{F_2(x)C(x)} + i\frac{G_1(x)}{G_2(x)} = \frac{F(x)}{F_2(x)C(x)} + i\frac{G_1(x)}{G_2(x)}.$$

$F(x)$ is a new symbol, written for $\sqrt{F_1(x)F_2(x)C^2(x)}$, and, as we shall

[3] Including as special cases real components which are simply rational functions, without the square root.

proceed to prove, it is the common numerator of all of the resistance and conductance expressions throughout the network.

Let us suppose now that the first branch, ia_1x, of the network is added in series. The admittance after its addition is

$$\frac{1}{\dfrac{F(x)}{F_2(x)C(x)} + i\dfrac{G_1(x)}{G_2(x)} + ia_1x} = \frac{F(x)}{F_2(x)C(x)\left[\dfrac{F_1(x)}{F_2(x)} + \left(\dfrac{G_1(x)}{G_2(x)} + a_1x\right)^2\right]}$$

$$- i\frac{F_2(x)\,C(x)\left(\dfrac{G_1(x)}{G_2(x)} + a_1x\right)}{F_2(x)\,C(x)\left[\dfrac{F_1(x)}{F_2(x)} + \left(\dfrac{G_1(x)}{G_2(x)} + a_1x\right)^2\right]}.$$

Upon remembering the way in which $C(x)$ was chosen we observe that the expressions in the denominators of the conductance and susceptance fraction and in the numerator of the susceptance fraction reduce to polynomials.

So far we have been able to show that the impedance of the load and the admittance of the network after one branch is added can be so expressed that (1) their imaginary components are rational functions, (2) the numerators of their real components are equal to $F(x)$, and (3) the denominators of their real components are polynomials. It is also possible, however, to show that if these statements hold for the impedance and admittance at *any* two consecutive junctions they will hold also at the next following junction. Referring to Fig. 3, let us

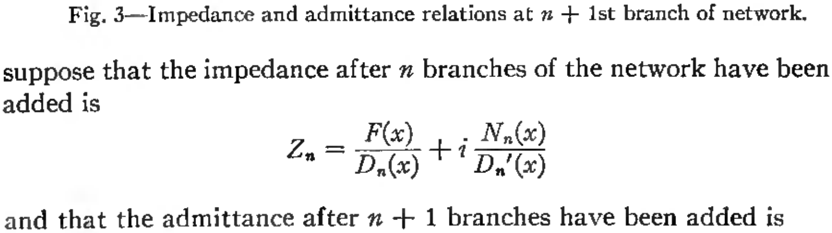


Fig. 3—Impedance and admittance relations at $n + 1$st branch of network.

suppose that the impedance after n branches of the network have been added is

$$Z_n = \frac{F(x)}{D_n(x)} + i\frac{N_n(x)}{D_n'(x)}$$

and that the admittance after $n + 1$ branches have been added is

$$Y_{n+1} = \frac{F(x)}{D_{n+1}(x)} + i\frac{N_{n+1}(x)}{D_{n+1}(x)}.$$

We wish to show that the impedance after the addition of the $n + 2nd$ branch is

$$Z_{n+2} = \frac{F(x)}{D_{n+2}(x)} + i\frac{N_{n+2}(x)}{D_{n+2}(x)}.$$

The various N's and D's, of course, represent polynomials. The denominator of the imaginary component of Z_n is accented, to indicate that it is not necessarily equal to the denominator of the real component. The denominators in the Y_{n+1} expression, however, have been given the same designation, since they are equal in the expression we have set up for the admittance at the terminals of the first network branch. This fact is not essential in the proof which follows, but its use somewhat simplifies the procedure. Direct mesh computation gives

$$Z_{n+2} = \frac{F(x)}{D_{n+1}\left[\dfrac{F^2}{D_{n+1}^2} + \dfrac{N_{n+1}^2}{D_{n+1}^2}\right] + a_{n+2}^2x^2D_{n+1} + 2a_{n+2}x}$$

$$- i\frac{[N_{n+1} + a_{n+2}xD_{n+1}]}{D_{n+1}\left[\dfrac{F^2}{D_{n+1}^2} + \dfrac{N_{n+1}^2}{D_{n+1}^2}\right] + a_{n+2}^2x^2D_{n+1} + 2a_{n+2}x}.$$

Since a_{n+2} is arbitrary, the resistance component will have the specified form only if

$$D_{n+1}\left[\frac{F^2}{D_{n+1}^2} + \frac{N_{n+1}^2}{D_{n+1}^2}\right]$$

is a polynomial in x. If this condition is satisfied the reactance expression can obviously be put in the required form.

In order to examine the denominator of the resistance expression more closely we state N_{n+1}, and D_{n+1} in terms of N_n, D_n', and D_n. Direct mesh computation, again, gives

$$Y_{n+1} = \frac{F(x)}{D_n\left[\dfrac{F^2}{D_n^2} + \dfrac{N_n^2}{D_n'^2}\right] + a_{n+1}^2x^2D_n + 2a_{n+1}xD_n\dfrac{N_n}{D_n'}}$$

$$- i\frac{\left[a_{n+1}x + \dfrac{N_n}{D_n'}\right]D_n}{D_n\left[\dfrac{F^2}{D_n^2} + \dfrac{N_n^2}{D_n'^2}\right] + a_{n+1}^2x^2D_n + 2a_{n+1}xD_n\dfrac{N_n}{D_n'}};$$

$$\therefore D_{n+1} = D_n\left[\frac{F^2}{D_n^2} + \frac{N_n^2}{D_n'^2}\right] + a_{n+1}^2x^2D_n + 2a_{n+1}xD_n\frac{N_n}{D_n'}$$

and

$$\frac{N_{n+1}}{D_{n+1}} = -\frac{\left[a_{n+1}x + \frac{N_n}{D'_n}\right] D_n}{D_n\left[\frac{F^2}{D_n^2} + \frac{N_n^2}{D_n'^2}\right] + a_{n+1}^2 x^2 D_n + 2a_{n+1}xD_n\frac{N_n}{D'_n}}$$

Substitution of these values for D_{n+1} and N_{n+1} reduces the expression for Z_{n+2} to

$$Z_{n+2} = \frac{F(x)}{D_n + a_{n+2}^2x^2D_{n+1} + 2a_{n+2}xN_{n+1}} - i\frac{N_{n+1} + a_{n+2}xD_{n+1}}{D_n + a_{n+2}^2x^2D_{n+1} + 2a_{n+2}xN_{n+1}}.$$

We have, however, assumed that D_n, D_{n+1}, and N_{n+1} were polynomials. The sums of the quantities constituting the numerator of the imaginary component of Z_{n+2} and the denominators of both components are therefore also polynomials, and, consequently, Z_{n+2} is written in the specified form.

The rest of the proof follows the usual argument from mathematical induction. In brief, we have established directly the fact that the formula holds when the network has no branches, or only one branch. Knowing that it holds for these two cases, we conclude from the above reasoning that it holds when there are two branches. If it is valid for one branch and two branches it must also be valid for three branches, and so on. Therefore the formula holds generally.

It will be observed that we have considered the admittance, rather than the impedance, when a series branch is added, and the impedance, rather than the admittance, when a shunt branch is added. Quite obviously the cases not considered are of little interest. If the analysis is stated in terms of impedance a final series branch contributes nothing to the resistance and can be considered as part of the reactance correcting network, while an analysis based upon admittances would similarly have no use for a final shunt branch except as a constituent of the susceptance correcting network. The general formula does hold, however, for these cases also. For example the addition of a series branch simply changes one rational function, representing the reactance at the terminals of the previous shunt branch into another rational function. The fact that the impedance at the terminals of the shunt branch falls into our general form is therefore sufficient to prove that the impedance after the series branch has been added can be written in this form also. This indicates, incidentally, that an alternative form of the proof we have been considering, based upon the impedance and admittance relation at a single junction, can be developed. Using the previous notation, the impedance Z_n will be in the required form if Z_{n+1} is in that form, and not otherwise. Instead of assuming that the impedance at one junction and the admittance at an adjacent junction can be fitted into the formula, therefore, it is sufficient to assume that both the impedance and admittance at a single junction satisfy the formula in order to show that the impedance and admittance at the next succeeding junction satisfy this formula also.

Application of General Analysis to Filter Impedance Correction

The reciprocal property of the impedances at the terminals of a reactive network indicates two possible methods of applying a ladder network of the sort we have been describing to the correction of wave filter impedances. We can either terminate the network by the filter impedance and adjust its parameters to match the line impedance, or we can consider that the load impedance of the network is a constant pure resistance, representing the line impedance, and attempt to produce a match at the filter terminals. These two methods of procedure lead to distinct results, since in one case the reactance or susceptance correcting branch adjoins the line, while in the other it adjoins the filter. Both are, however, admissible under the general mathematical specifications we have set up for the load impedance of the resistance or conductance controlling network and both lead to reasonably satisfactory impedance correction.

The fact that a constant pure resistance is an admissible load impedance for the ladder network is easily established by inspection. The rational function $G_1(x)/G_2(x)$, representing the imaginary component, reduces to zero, of course, while the rational function $F_1(x)/F_2(x)$, whose square root represents the real component becomes a constant. A filter image impedance within transmission bands is similarly a pure resistance. As a function of frequency it may be defined as the geometric mean of the open and short-circuit impedances. An open or short-circuit filter, whatever its configuration is, however, simply a network of pure reactances. The open and short-circuit impedances are therefore rational functions of frequency and the image impedance they define falls within the scope of the mathematical specification we have set up for the load impedance of the correcting network.

Terminating Networks of the First Type

While both of these methods of approaching the problem lead to satisfactory impedance correction, other considerations to be discussed later recommend that one in which the filter is taken to be the load impedance for most designs. This approach will therefore be considered first and in greatest detail. We will, moreover, limit ourselves to image impedances of the "constant-k" type. Practical filter designs of course are usually composite structures containing several types of sections. The image impedances at the junctions between the sections are however, nearly always of the "constant-k" type and our restriction to image impedances belonging to this class does not, therefore, seriously limit the field of application for the network.

Notation

The image impedance of a mid-series terminated "constant-k" filter is usually written as

$$Z_0\sqrt{1+\frac{Z_{1k}}{4Z_{2k}}};$$

that of a mid-shunt terminated filter as

$$\frac{Z_0}{\sqrt{1+\frac{Z_{1k}}{4Z_{2k}}}},$$

where Z_{1k} and Z_{2k} represent in each case the series and shunt impedances of the "constant-k" filter, and $Z_0(=\sqrt{Z_{1k}Z_{2k}})$ is a constant which can be chosen arbitrarily to fix the impedance level of the circuit.

We will find it convenient to represent the way in which the various branches vary with frequency by a new quantity x, defined by the relation

$$\frac{Z_{1k}}{2}=iZ_0x.$$

In a low pass filter, for example, $x=f/f_c$, in a high-pass filter $x=f_c/f$, and in a band-pass filter

$$x=\frac{\frac{f}{f_m}-\frac{f_m}{f}}{\sqrt{\frac{f_2}{f_1}}-\sqrt{\frac{f_1}{f_2}}}.$$

Upon making use of the relation $Z_{1k}Z_{2k}=Z_0^2$ the formulæ for mid-series and mid-shunt "constant-k" image impedances can be written as $Z_0\sqrt{1-x^2}$ and $Z_0/\sqrt{1-x^2}$.[4]

This method of representing the image impedances suggests that

[4] In terms of the usual filter notation this $x=\sqrt{-U_k}$.

the configuration of the resistance or conductance controlling network be so chosen that the impedances of its series branches and the admittances of its shunt branches are proportional to x. In other words the series and shunt branches of the correcting network should be similar physically to those of the "constant-k" filter. The complete network is then that shown in Fig. 4. It is similar to that of Fig. 2

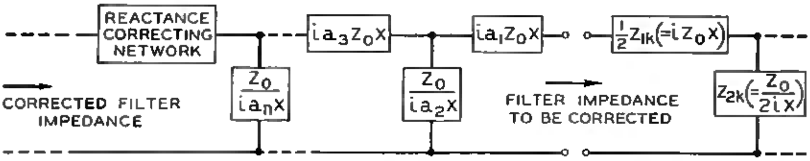


Fig. 4—Generalized schematic of first or "direct" type of filter terminations.

except that the explicit introduction of the factor Z_0 into the expressions for the series and shunt branches reduces the a's to constants of proportionality which can be fixed, once for all, for all "constant-k" filters. Following the analogy of ordinary filter structures it will be assumed that the first branch of the network is in series when the filter proper is mid-series terminated, and vice versa. It is then easily shown that the preceding general formula for the resistance[5] of the system reduces, both for mid-series and mid-shunt terminated filters, to

$$R = \frac{Z_0\sqrt{1-x^2}}{1 + A_1x^2 + A_2x^4 + \ . \ . \ . \ A_nx^{2n}},$$

when n is the number of branches in the network. It will be observed that odd powers of x are missing.

The possibilities of manipulating this expression to secure desirable resistance characteristics are obviously determined by the number, n, of variable terms in the denominator of the expression. Since n is, however, also equal to the number of branches of the resistance or conductance controlling network, and therefore determines both the cost of this network and the extent to which the resistance or conductance can be made to approximate a given curve, it offers a convenient basis for differentiating between the various structures. The simplest cases, and the only ones of practical importance in contemporary filter design, are those for which $n = 1$, 2, or 3. They are illustrated in Fig. 5 and will be taken up in order. Our first step will be the establishment of the algebraic relations between the element values $a_1 \ . \ . \ . \ a_n$ and the parameters $A_1 \ . \ . \ . \ A_n$ for each of these three cases.

[5] Assuming that the final branch, $1/ia_nx$ is in shunt as in Fig. 4. When the analysis is stated in terms admittances the results are precisely similar, except for an obvious change from Z_0 to $1/Z_0$.

The analyses are stated in terms of conductance and susceptance, since in this form they are most conveniently applied to the impedance correction of systems of parallel filters, which constitute a large proportion of practical cases. The formulæ and curves can, however, be used directly in analyses stated in terms of impedance if we merely replace conductance and susceptance by resistance and reactance and write Z_0 in the numerator rather than in the denominator whenever it

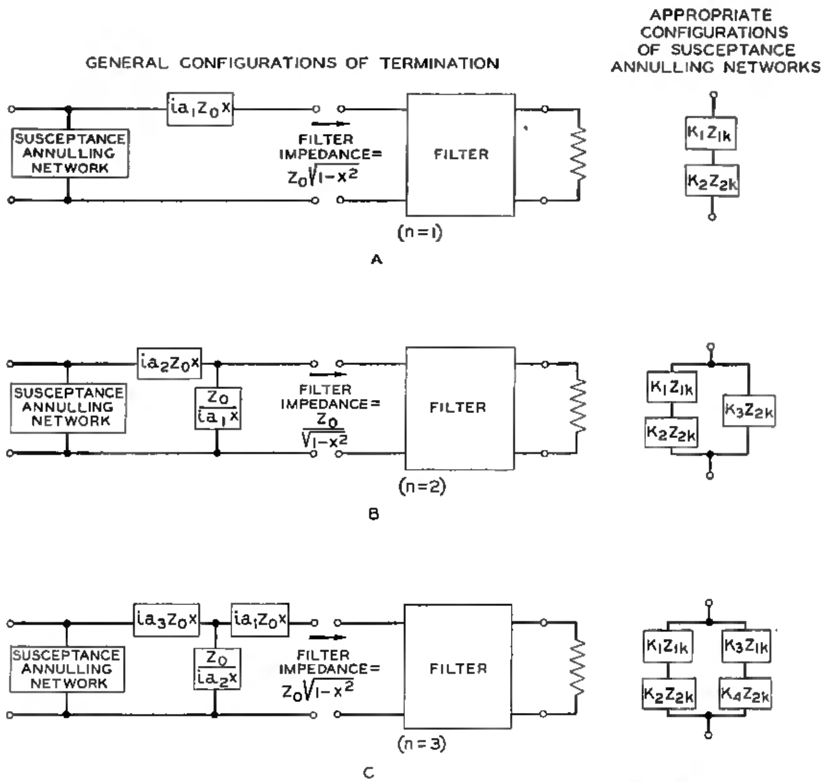


Fig. 5—Configurations of 1, 2 and 3 branch terminations.

appears. After the relations between $a_1 \ldots a_n$ and $A_1 \ldots A_n$ have been determined we shall proceed to a discussion of methods of choosing values of $A_1 \ldots A_n$ giving a suitable resistance or conductance characteristic. The final steps are the computation of the element values of the network from these values of the polynomial coefficients, the calculation of the resulting reactance or susceptance characteristic and the design of a final branch giving the complete structure the desired reactance or susceptance characteristic.

Analytical Relations between Polynomial Coefficients and Element Values

Case I—$n = 1$.

The general analysis shows that the conductance of the system must be expressible in the form

$$G = \frac{1}{Z_0}\frac{\sqrt{1 - x^2}}{1 + A_1 x^2}.$$

A direct mesh computation of the network of Fig. 5-*a* gives

$$G = \frac{1}{Z_0}\frac{\sqrt{1 - x^2}}{1 - (1 - a_1^2)x^2}.$$

From which, by comparison of coefficients,

$$A_1 = -(1 - a_1^2)$$

or

$$a_1 = \sqrt{1 + A_1}.$$

The susceptance characteristic is given by

$$B = -\frac{1}{Z_0}\frac{a_1 x}{1 + A_1 x^2}.$$

It can be annulled exactly by the reactance

$$iX = i\,\frac{1 - a_1^2}{a_1} Z_0 x + \frac{Z_0}{i a_1 x} = \left(\frac{1 - a_1^2}{2a_1}\right) Z_{1k} + \frac{2}{a_1} Z_{2k}$$

where Z_{1k} and Z_{2k} are, as before, the series and shunt impedances of the "constant-*k*" filter.

If the conductance and susceptance controlling portions of the network are combined the resulting structure is identical with a half section of the conventional "m-derived" type. We have merely to replace a_1 by m. Single branch conductance controlling networks therefore contribute nothing new to filter impedance correction. Multiple branch networks, which can be considered, if one pleases, as natural extensions of the "m-derived" scheme, must be looked to for the solution of impedance problems for which standard sections are inadequate.

Case II—$n = 2$.

A direct computation of the network shown in Fig. 5-*b* gives

$$G = \frac{1}{Z_0}\frac{\sqrt{1 - x^2}}{1 + (a_2^2 - 2a_1a_2)x^2 + a_2^2(a_1^2 - 1)x^4},$$

whence

$$A_1 = a_2^2 - 2a_1a_2,$$
$$A_2 = a_2^2(a_1^2 - 1),$$

$$a_1 = \frac{1 \pm \sqrt{1 + A_1 + A_2}}{\sqrt{(1 \pm \sqrt{1 + A_1 + A_2})^2 - A_2}},$$

$$a_2 = \sqrt{(1 \pm \sqrt{1 + A_1 + A_2})^2 - A_2}.$$

The upper of the alternative signs usually gives the better reactance characteristic.

The susceptance of this network is

$$B = -a_2x \frac{\left(1 - \frac{a_1}{a_2}\right) - (1 - a_1^2)\,x^2}{1 + A_1x^2 + A_2x^4}.$$

Case III—$n = 3$

The general conductance expression is

$$G = \frac{1}{Z_0} \frac{\sqrt{1 - x^2}}{1 + A_1x^2 + A_2x^4 + A_3x^6},$$

where

$$A_1 = a_1^2 + 2a_1a_3 + a_3^2 - 2a_2a_3 - 1,$$
$$A_2 = a_2^2a_3^2 + 2a_2a_3 - 2a_1^2a_2a_3 - 2a_1a_2a_3^2,$$
$$A_3 = a_1^2a_2^2a_3^2 - a_2^2a_3^2.$$

These equations can be reduced to

$$a_1 + a_3 - a_1a_2a_3 = \pm \sqrt{1 + A_1 + A_2 + A_3},$$
$$a_1 + a_3 = \sqrt{1 + A_1 + 2a_2a_3},$$
$$a_1a_2a_3 = \sqrt{A_3 + a_2^2a_3^2},$$

from which

$$\sqrt{1 + A_1 + 2a_2a_3} - \sqrt{A_3 + a_2^2a_3^2} = \pm \sqrt{1 + A_1 + A_2 + A_3}.$$

Upon examining the form of the radicals on the left we see that a_2a_3 is determined by the intersection of a parabola and a hyperbola.

Once a_2a_3 are known the individual values of a_1, a_2, and a_3 can be found directly from the previous equations. The two radicals on the left side of the equation must be taken as positive in order to secure positive elements, which is the same as saying that the two conic sections must intersect in the first quadrant. The square root on the right hand side may be taken either as positive or negative, the susceptance characteristic obtained with the negative sign being usually preferable.

It is also possible to eliminate two of the a's directly, obtaining the equation

$$\begin{aligned}[A_2^2 - 4A_2A_3 - 4A_3]a_1^4 + 8A_3\sqrt{1 + A_1 + A_2 + A_3}\,a_1^3 \\ - [2A_1^2 + 2A_2A_3 - 4A_1A_3]a_1^2 - 8A_3\sqrt{1 + A_1 + A_2 + A_3}\,a_1 \\ + [(A_2 + A_3)^2 + 4A_3] = 0,\end{aligned}$$

which can be solved by standard methods. The former method is shorter, however.

The susceptance is given by

$$B = -x\frac{B_0 + B_1x^2 + B_2x^4}{1 + A_1x^2 + A_2x^4 + A_3x^6},$$

where

$$B_0 = a_1 + a_3 - a_2,$$

$$B_1 = a_2 + a_3a_2^2 - a_1^2a_2 - 2a_1a_2a_3,$$

$$B_2 = a_1^2a_2^2a_3 - a_2^2a_3.$$

Methods of Choosing Power Series Coefficients

Having developed the relations between the power series coefficients and the network elements we are now ready to consider methods of choosing the parameters to fit given impedance requirements. Upon rewriting our equation for the real component of the network admittance in the form

$$1 + A_1x^2 + A_2x^4 + \cdots A_nx^{2n} = \frac{\sqrt{1 - x^2}}{Z_0G}$$

we see that the problem reduces to the approximation of the ratio of $\frac{1}{Z_0}\sqrt{1 - x^2}$ to the desired conductance G, both of which are known, by means of the polynomial $1 + A_1x^2 + \ldots + A_nx^{2n}$. In most practical designs the desired filter impedance will be a constant resistance. It is then convenient to rewrite the equation as

$$\frac{Z_0}{R_0}(1 + A_1 x^2 + \cdots A_n x^{2n}) = \sqrt{1 - x^2},$$

where R_0 denotes the desired constant resistance. The problem thus becomes that of simulating $\sqrt{1 - x^2}$ in the range $0 < x < 1$ by means of a polynomial in x^2 of degree n, and if we assume that the parameter Z_0 can be chosen arbitrarily the polynomial is completely unrestricted, since the constant term as well as the coefficients of the various powers of x can be taken at pleasure.

There are several ways of proceeding from this point. The simplest makes use of the binomial theorem. Upon expanding $\sqrt{1 - x^2}$ with the help of this theorem we reach the relation

$$\frac{Z_0}{R_0}(1 + A_1 x^2 + \cdots A_n x^{2n}) = 1 - \frac{1}{2}x^2 - \frac{1}{8}x^4 - \frac{1}{16}x^6 \cdots$$

Equating corresponding powers of x gives

$$\begin{aligned} Z_0 &= R_0, \\ A_1 &= -1/2, \\ A_2 &= -1/8, \\ A_3 &= -1/16, \\ &\cdots\cdots \\ &\cdots\cdots \end{aligned}$$

Using n branches in the conductance controlling network it is possible to take the first n terms of the binomial expansion into account. The elements corresponding to these values of A_1, A_2, etc. can of course be found by the equations derived previously. The results are summarized in the following table.

TABLE I

Number of Branches	A_1	A_2	A_3	a_1	a_2	a_3
1	− 0.5000	0	0	0.7071	0	0
2	− 0.5000	− 0.1250	0	0.97679	1.6507	0
3	− 0.5000	− 0.1250	− 0.0625	1.00308	1.96227	1.62715

The conductance characteristics corresponding to these choices of parameters are shown on Fig. 6. The curve $n = 0$, which corresponds to the "constant-k" type image impedance, has also been added for comparison. It will be seen from the curves that these values of the coefficients $A_1 \ldots A_n$ give very good approximations for small values of x, but inferior ones for values near unity. It is preferable in most designs to sacrifice something at the lower end of the characteristic in order to secure better performance in the higher range.

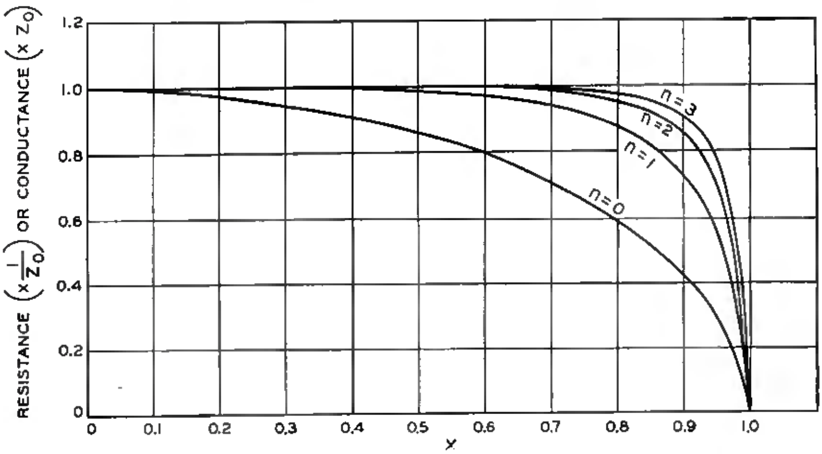


Fig. 6—Resistance and conductance characteristics secured from the binomial expansion.

The advantage of an approximation distributed over the band is gained by an expansion in terms of Legendrian harmonics. These functions are discussed in standard reference books, such as Byerly "Fourier Series and Spherical Harmonics" or Whittaker and Watson "Modern Analysis." It is important to mention here, however, that they are simply polynomials. Any polynomial such as $\frac{Z_0}{R_0}(1 + A_1x^2 + \ldots A_nx^{2n})$ can be broken up into a linear combination of even ordered harmonics, and, conversely, any linear combination of even ordered harmonics can be reduced to the form $\frac{Z_0}{R_0}(1 + A_1x^2 + \ldots A_nx^{2n})$. It is therefore easy to convert an expansion in terms of even harmonics into a power series of the sort with which we are directly concerned. The property of these functions of most interest here is the fact that, for an expansion of any given degree, they give the best "least squares" approximation to the desired function. In the range between $x = 0$ and $x = 1$, therefore, the approximation they furnish is much better for most purposes than that given by the binomial theorem. The expansion of $\sqrt{1 - x^2}$ in terms of Legendrian harmonics is given on p. 184 of Byerly as

$$\sqrt{1-x^2} = \frac{\pi}{2}\left[\frac{1}{2}P_0(x) - 5\left(\frac{1}{4}\right)\left(\frac{1}{2}\right)^2 P_2(x) - 9\left(\frac{3}{6}\right)\left(\frac{1}{2.4}\right)^2 P_4(x) - 13\left(\frac{5}{8}\right)\left(\frac{1.3}{2.4.6}\right)^2 P_6(x) + \ldots\right].$$

Upon replacing the harmonics by their values in terms of x,—

$$P_0(x) = 1,$$

$$P_2(x) = \frac{1}{2}(3x^2 - 1),$$

$$P_4(x) = \frac{1}{8}(35x^4 - 30x^2 + 3),$$

$$P_6(x) = \frac{1}{16}(231x^4 - 315x^4 + 105x^2 - 3),$$

and sorting out the various powers of x, values of the coefficients $A_1 \ldots A_n$ are secured, and from these the actual element values are found by means of formulæ developed previously. The following table summarizes the results

TABLE II

Number of Branches	$\frac{R_0}{Z_0}$	K_0	K_1	K_2	K_3	
0	1.273	0.7855	0	0	0	
1	0.9699	0.7855	− 0.4909	0	0	
2	1.011	0.7855	− 0.4909	− 0.1105	0	
3	0.9948	0.7855	− 0.4909	− 0.1105	− 0.04986	
Number of Branches	A_1	A_2	A_3	a_1	a_2	a_3
0	0	0	0	0	0	0
1	− 0.7142	0	0	0.5546	0	0
2	− 0.3236	− 0.4884	0	0.8986	1.593	0
3	− 0.0461	+ 0.4958	− .7162	0.9597	1.924	1.565

The quantities $K_0 \ldots K_3$ are the numerical coefficients of the corresponding harmonics. It will be observed that with this method of determining the network parameters Z_0 is not quite equal to R_0. When the analysis is based upon impedances instead of admittances the ratio R_0/Z_0 should be replaced by Z_0/R_0. The conductance characteristics secured by this process are shown in Fig. 7.

It is, of course, always possible to dispense with these general methods entirely and make an empirical determination of the design parameters. The particular requirements of specific design projects

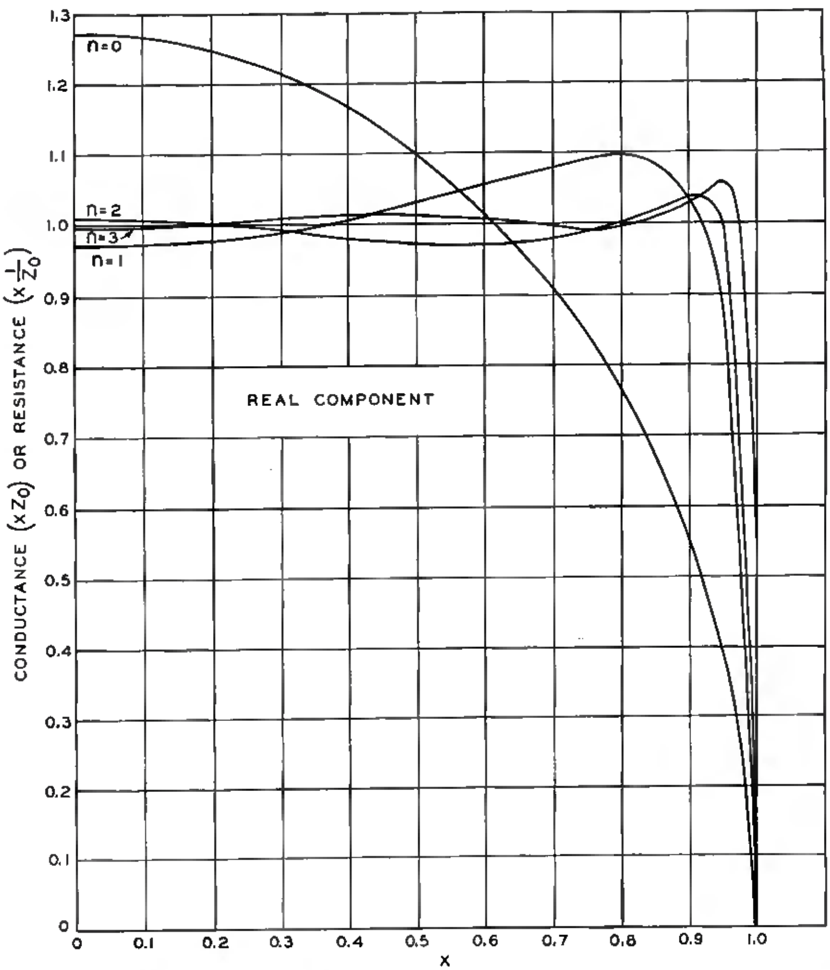


Fig. 7—Resistance and conductance characteristics secured from expansion in terms of Legendrian harmonics.

are thereby given the fullest recognition. This method was used in constructing the sections described in the accompanying paper. Even when the empirical method is adopted, however, the networks determined by the general expansions, particularly that in terms of Legendrian harmonics, should be valuable as starting points.

In most designs it is desirable to make the maximum departure from the ideal characteristic within the operating range as small as possible. A method of doing this for the 2-branch networks has been developed. The method assumes that the Z_0 of the filter has been taken equal to the terminating impedance, which assures a correct conductance at the point $x = 0$. The manipulation of the parameters A_1 and A_2 allows us to secure the desired value of conductance at two additional points. The result is a two looped characteristic, similar, if we make allowance for the difference in the assumptions regarding Z_0, to that already determined for this network by means of the Legendrian expansion. The requirement that the maximum departure from the ideal within the operating range be a minimum is equivalent to saying that the amplitudes of the downward and upward loops must be equal. It can be shown that a 2-branch conductance network will satisfy this condition if

$$\frac{-\frac{27}{16}A_2}{1 + A_1 + A_2} = (A_1^2 - 4A_2)^2.$$

In view of the relations which have been developed between A_1, A_2,

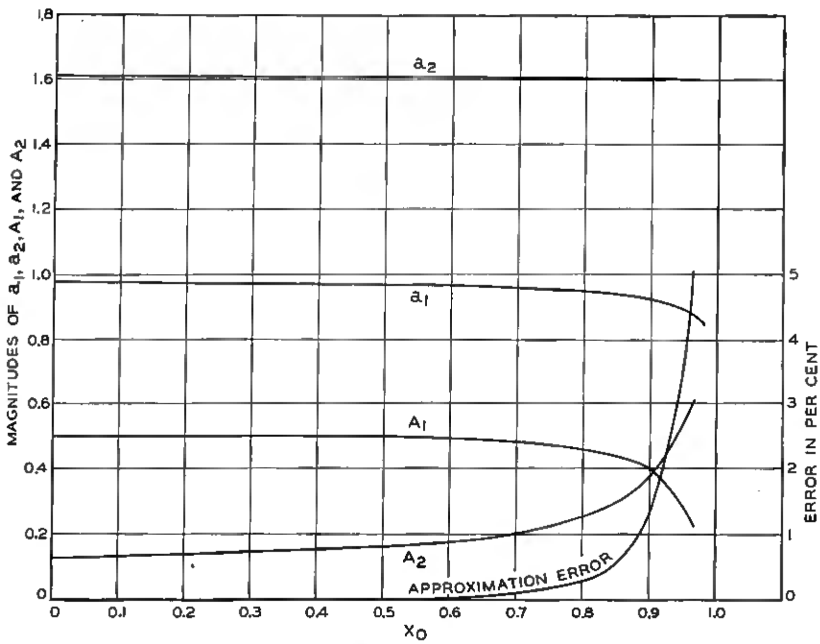


Fig. 8—Design chart for 2-branch termination.

and a_1, a_2 this condition can also be written in the form

$$\frac{27}{16}(1 - a_1^2) = a_2^2[4(1 - a_1a_2)^2 + a_2^2(1 - a_1a_2)]^2.$$

A second condition upon these quantities is found by specifying the range within which the impedance is to remain as flat as possible. The results of computations to determine this relationship are given in Fig. 8. x_0 in this diagram signifies the highest value of x in the operating range. Fig. 8 also gives the maximum departure of the conductance characteristic from its ideal value as a function of x_0. Numerical data taken from these curves should of course be confirmed by the equations given herewith before they are used to specify element values.

Susceptance Correcting Networks [6]

Once the conductance controlling portion of the network has been determined by one or another of these methods our general procedure calls for the computation of the susceptance characteristic it furnishes and the design of a final shunting reactance network which will annul

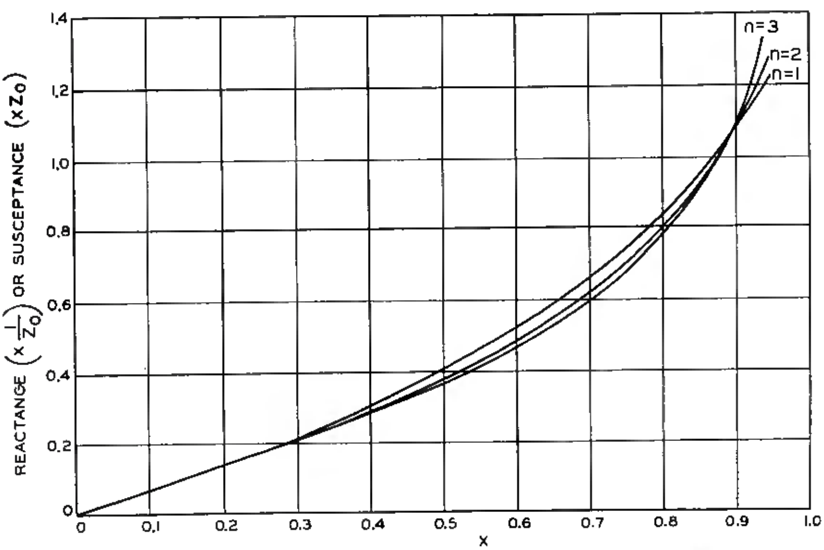


Fig. 9—Reactance and susceptance characteristics secured from binomial expansion.

[6] This section gives only a general description of the characteristics required of the susceptance correcting networks and the configurations which have been found appropriate for them. The design of these networks may be conveniently approached by means of the formulæ contained in R. M. Foster's article "A Reactance Theorem," in the Oct. 1924 issue of this *Journal.*

this susceptance to a suitable approximation. Fortunately the characteristics required of this network are of a type which can readily be obtained with physically realizable elements. The curves of Figs. 9 and 10 represent the susceptance characteristics required for the

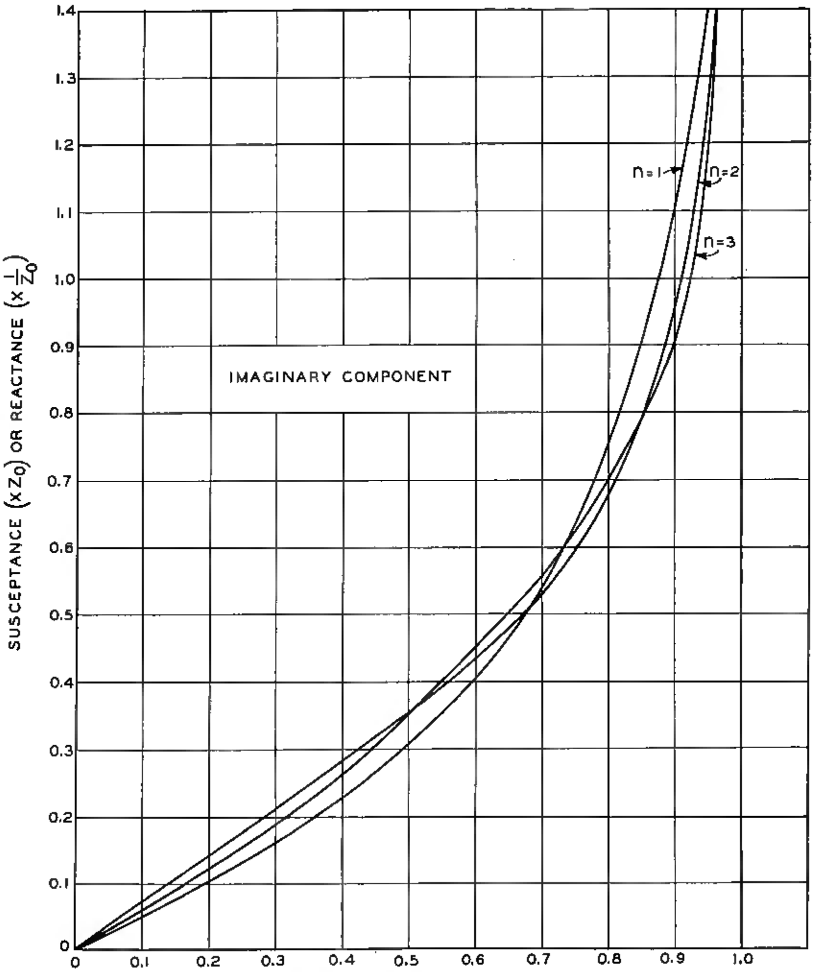


Fig. 10—Reactance and susceptance characteristics secured from expansion in terms of Legendrian harmonics.

Legendrian and binomial expansion networks. Empirically determined networks give very similar results. The general configuration of appropriate susceptance correcting networks can be determined from an inspection of these curves. For example, if we assume that a low pass filter is in question, which means that the variable "x"

is proportional to frequency, the desired susceptance curves will be recognized as being approximately those which would be obtained from tuned circuits resonating slightly beyond the cutoff. Since a tuned circuit can be considered as being a series combination of the series and shunt impedance of the "constant-k" filter, any such correcting network designed for a low-pass filter can be adapted to

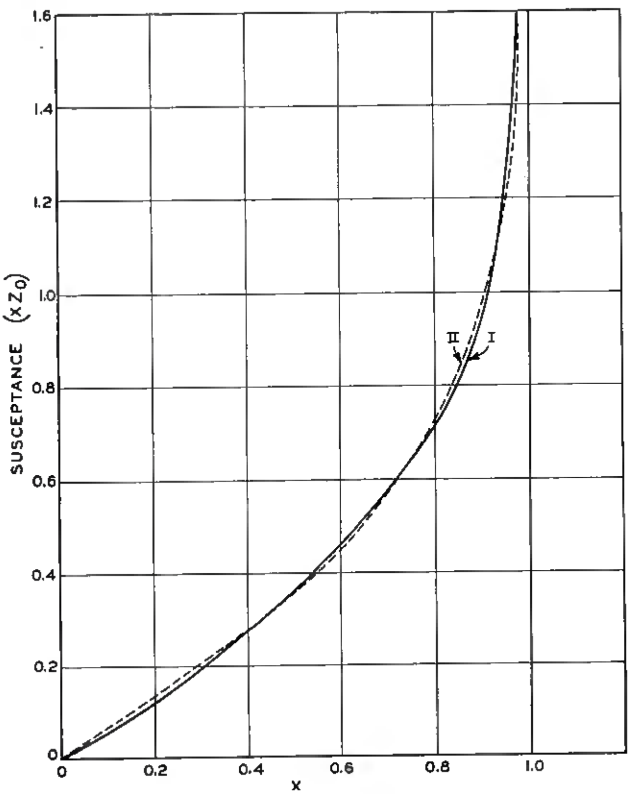


Fig. 11—Susceptance correction of a 3-branch termination.
I—Desired susceptance.
II—Susceptance actually obtained.

another type of "constant-k" structure by replacing inductances and capacities by the homologous impedances of the other filter.

This simple combination of series and shunt impedances is, as we have previously seen, capable of giving exact susceptance correction when the conductance controlling network contains only one branch, but it is not, in general, sufficient for 2 and 3 branch networks. Indeed, no physically realizable reactive network will cancel the susceptance furnished by these more complicated structures exactly. Close approximations however can be obtained by modifying the "tuned circuit" characteristic slightly through the introduction of extra elements. Suitable configurations for 2 and 3 branch networks have already been given in Fig. 5. They should furnish susceptance characteristics at least as good as the corresponding conductance characteristics. An example of the susceptance correction of a three branch network, using the configuration of Fig. 5-c, is shown in Fig. 11. Curve I represents the ideal susceptance characteristic, Curve II that actually obtained.

Impedance Correction of Paralleled Filters

An interesting modification of the process of susceptance correction occurs when a number of filters are to be connected in parallel. Since the impedance of an attenuating filter is almost a pure reactance the conductance component of a system of parallel filters at a given frequency is furnished almost entirely by the filter in whose transmission band that frequency lies. If the system as a whole is to have the correct conductance throughout each transmission band, therefore, every filter must be given the conductance controlling network which would be appropriate if it were operating alone. While the process of conductance correction is thus exactly the same for multipled and individual filters, the process of susceptance correction of paralleled filters must be modified somewhat to take account of the susceptance component furnished by the attenuating filters. A single susceptance network will serve for the whole system. We have merely to compute the susceptance characteristics furnished by the various filters terminated in their conductance controlling networks and annul them throughout every transmission band by a two terminal network in parallel with the system as a whole. An example of the application of the method to a pair of parallel complementary filters having 2 branch conductance controlling networks is given by Fig. 12. Curve I in this diagram represents the susceptance of the transmitting filter, Curves II the susceptance of the attenuating filter for several different choices of its cutoff frequency, Curves III the susceptances of the corresponding auxiliary networks, and Curves IV the net result. A series combination of the series and shunt impedances of either filter[7] resonating at the geometric mean of the cutoff frequencies was chosen for the

[7] Since the filters are complementary the series impedance of one is similar to the shunt impedance of the other, and vice versa. By choosing the resonance frequency of the auxiliary network symmetrically with respect to the two filters, as we have done, all of the susceptance relations become symmetrical, and the network functions as well for one filter as it does for the other.

auxiliary network. By using two resonant arms with closely adjacent resonance frequencies still better susceptance correction could have been secured.

Filters which must operate in parallel are usually given x-terminations. Since an x-termination can be thought of as being a one element

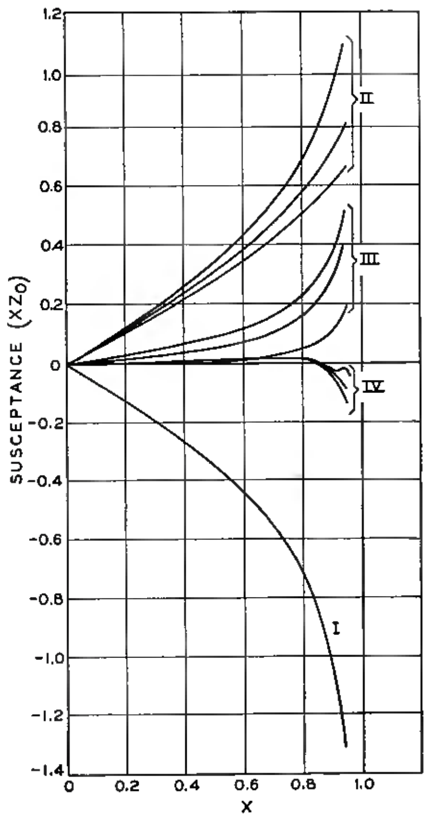


Fig. 12—Susceptance relations at the line terminals of a pair of parallel complementary filters having 2-branch conductance controlling networks.

conductance controlling network the method we are discussing can be applied here also. It is interesting to note that the introduction of an auxiliary susceptance controlling network considerably improves the performance even of this well known circuit. The susceptance relations at the line terminals of a pair of parallel complementary x-terminated filters are shown in Fig. 13, the arrangement of the curves being similar to that of Fig. 12. The improvement can be estimated from the magnitude of the auxiliary susceptance.

The auxiliary network improves the susceptance of parallel band pass filters even more than it does that of complementary filters. Curve I of Fig. 14 represents the susceptance of a typical uncorrected set of band pass filters. The first step in the improvement of this characteristic is due to Mr. R. H. Mills, who suggested that networks whose impedances resemble that of filters above and below the actual set of bands be added to the system. This reduces the susceptance

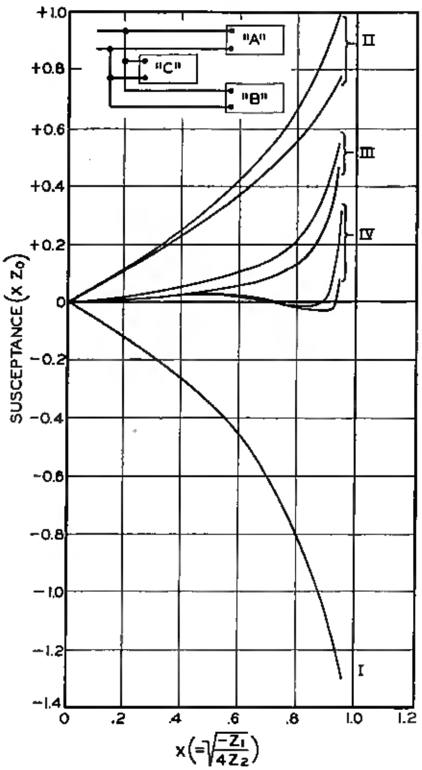


Fig. 13—Susceptance relations at the line terminals of a pair of parallel complementary x-terminated filters.

to the level shown by Curve II. Curve III gives the completely corrected characteristic. The auxiliary susceptance correcting network consists of a number of tuned circuits in parallel, one resonating between each pair of successive bands, together with one resonating above the topmost band and one resonating below the lowest band. The insertion of the auxiliary network has the further advantage that it produces peaks of attenuation near the cutoffs of the filters, thus enhancing their selectivity.

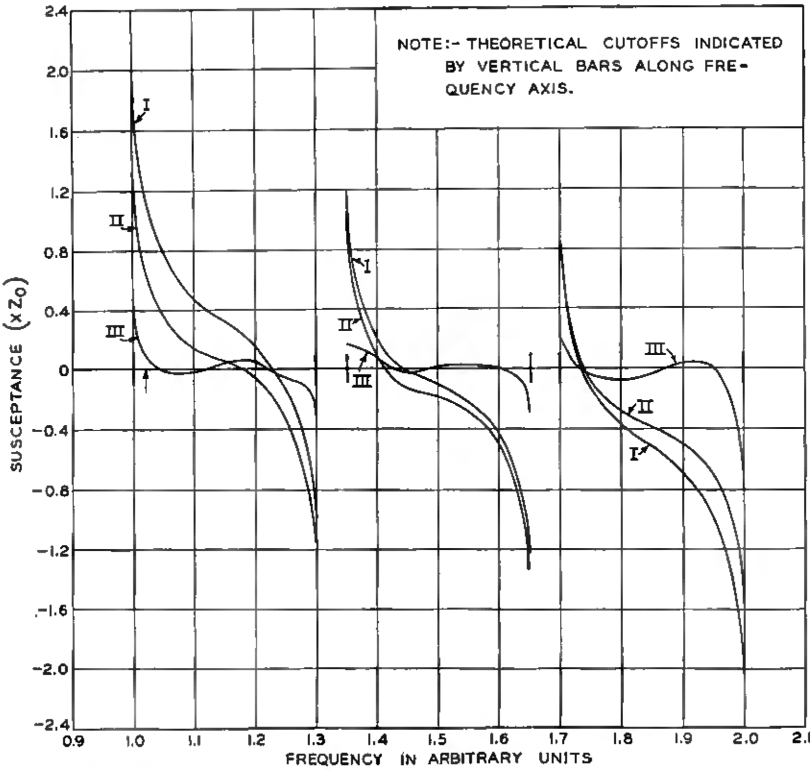


Fig. 14—Susceptance correction of a set of parallel x-terminated band-pass filters.

I—Uncorrected susceptance.
II—Susceptance after the addition of a simple auxiliary network.
III—Susceptance after the addition of a more elaborate auxiliary network.

Reverse Method of Designing Terminating Sections

Hitherto we have assumed that the load impedance of the terminating network was the filter image impedance, and our procedure has consisted essentially in determining an adjustment of the network parameters which would make its input impedance a constant pure resistance. As we have already seen, however, it is equally legitimate to assume that the network is terminated in the line resistance, and determine parameter values which will produce a match between its impedance and that of the filter. This assumption leads to the circuit arrangement shown in Fig. 15.

Upon examining what happens to the general expression for the resistance of the network when the load impedance reduces to the constant pure resistance, R_0, we easily find that it turns out to be

$$R = \frac{R_0}{1 + A_1x^2 + A_2x^4 + \ldots + A_nx^{2n}},$$

where n is the number of branches in the network. Odd powers of x are missing, just as they were when the network was terminated in a filter impedance.

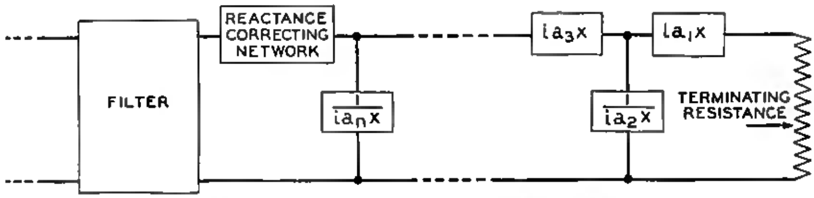


Fig. 15—Generalized schematic of second or "reverse" type of filter terminations.

Our problem consists in matching this expression to the filter impedance, $Z_0\sqrt{1 - x^2}$. Upon assuming that $R_0 = Z_0$, for simplicity, we see that it reduces to the selection of values of $A_1 \ldots A_n$ which will secure approximate satisfaction of the equation

$$1 + A_1x^2 + \ldots + A_nx^{2n} = \frac{1}{\sqrt{1 - x^2}}$$

Two empirical [8] choices of these parameters have been made, one

[8] Our previous methods of approximation, in terms of Taylor's series and Legendrian harmonics, are of course available here also. In addition, if we rewrite the expression as

$$\sqrt{1 - x^2}\,\frac{Z_0}{R_0}(1 + A_1x^2 + \ldots + A_nx^{2n}) = 1$$

the left hand side appears as a linear combination of the associated Legendrian functions $P_1'(x)$, $P_3'(x)$, $\ldots$, defined by the general formula

$$P_n'(x) = \sqrt{1 - x^2}\,\frac{d}{dx}P_n(x),$$

where $P_n(x)$ is the usual Legendrian function. The problem can therefore be considered as that of approximating unity by a series of the associated functions. These methods of approach differ chiefly in the relative weights which they ascribe to various portions of the frequency band. Judged by this criterion neither of the first two methods is very satisfactory for practical applications. The Taylor's series expansion, of course, is best in the neighborhood of $x = 0$. The "least squares" property of the ordinary Legendrian functions, on the other hand, tends to produce rough equality in the numerical values of the departures from the desired function in various portions of the frequency band. From the engineering standpoint, however, it is the percentage departure from the desired impedance, and not the numerical departure, which is of interest. This type of approximation therefore leads to a relative over-emphasis of the region near $x = 1$, where the desired function $1/\sqrt{1 - x^2}$ is large. The approach by means of the associated functions, however, avoids this objection, since the approximated function is in this case a constant, and leads to characteristics substantially as good as those obtained by means of the empirically determined parameters discussed in the text.

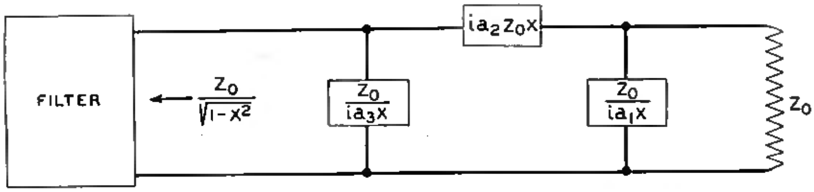


Fig. 16—A 2-branch termination of the "reverse" type.

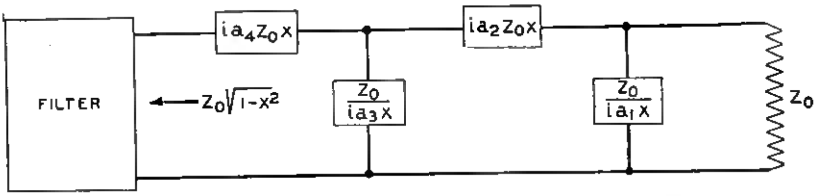


Fig. 17—A 3-branch termination of the "reverse" type.

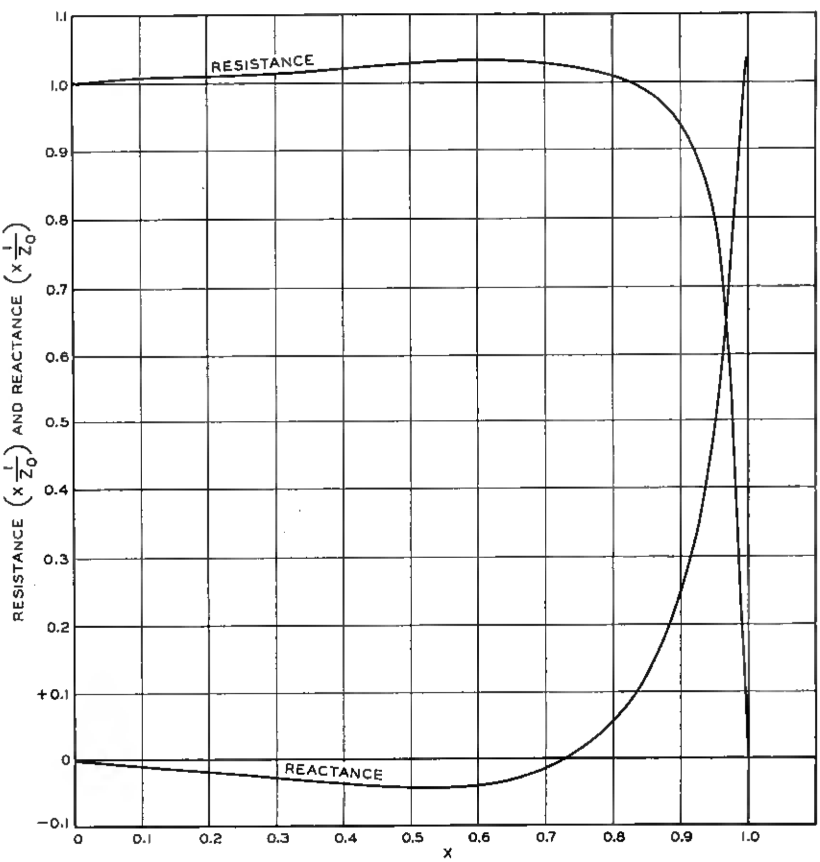


Fig. 18—Impedance characteristic secured from the network of Fig. 16.

when the network contained two branches, and the other when it contained three. In both instances the appropriate reactance or susceptance annulling networks were found to be simple arms, similar to the series or shunt branches of the remainder of the termination in physical configuration. The complete networks are shown in Figs. 16

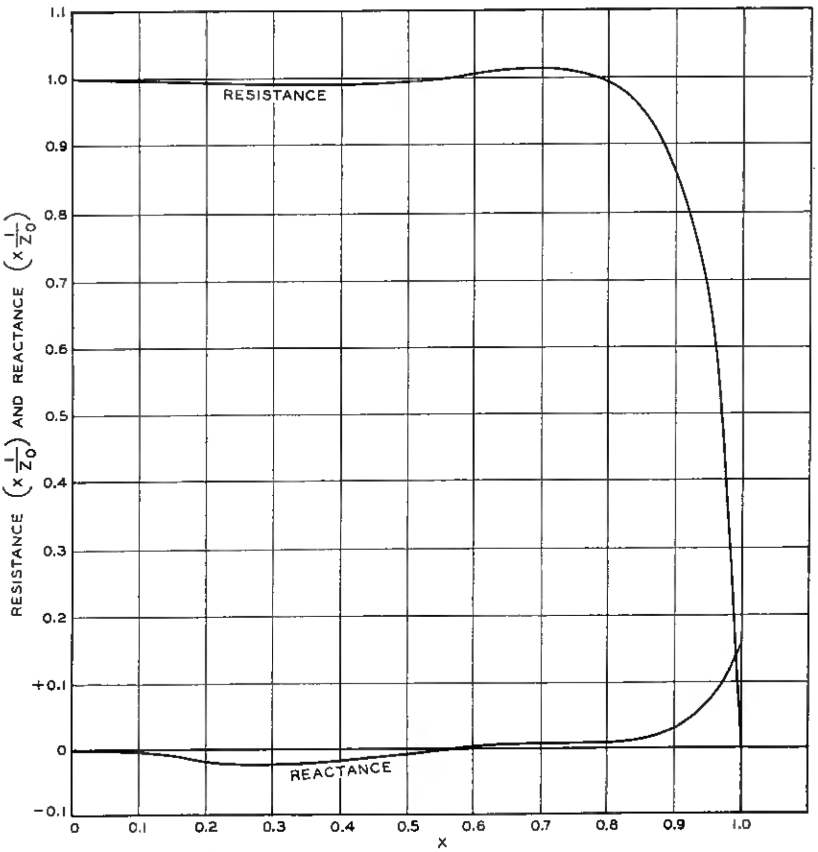


Fig. 19—Impedance characteristic secured from the network of Fig. 17.

and 17, where the final branches, Z_0/ia_3x in Fig. 16, and ia_4Z_0x in Fig. 17 are the susceptance or reactance annulling networks. The values of the various parameters are given in the following table.

TABLE III

n	A_1	A_2	A_3	a_1	a_2	a_3	a_4
2	+ 0.0505	+ 1.6508	0	0.7973	1.6186	0.904	0
3	+ 0.9114	− 1.8488	+ 3.2823	0.6733	1.466	1.835	0.925

If a perfect match were secured at the filter terminals then, by the reciprocity principle, a perfect match should be secured at the line terminals also. In order to evaluate the performance of the networks, therefore, the impedances they present to the line were computed. The results are shown in Figs. 18 and 19.

Comparison of Direct and Reverse Networks

At first glance the curves of Figs. 18 and 19 seem to show that while networks of the reverse type produce a good impedance match over a moderate fraction of band they will be much less successful than the structures previously described at frequencies very near the cutoff. This apparent advantage in favor of the networks first described is discounted considerably however by the economy of elements resulting from the relative simplicity of the reactance or susceptance controlling networks used with terminations of the second type. If we adopt as our standard in comparing the two types of networks the total number of elements each requires, rather than the number of branches they contain, the advantage of networks of the first type becomes much less impressive, if it does not actually disappear. More important considerations recommending the first type of terminations in preference to the second for most practical designs appear to be the greater ease with which they can be designed to meet a given reflection coefficient requirement, resulting from the relatively smaller number of branches they contain, the greater ease with which they can be adapted to filters which must operate in parallel, and the fact that the attenuation they contribute to the total filter suppression is usually more useful than that furnished by terminations of the second type.

Under certain circumstances, however, the second type of terminating sections have a definite advantage over the others. When a filter operates in conjunction with a modulating device a high modulator efficiency with low distortion demands that the impedance of the filter to the untransmitted side band be low (or high) and nearly constant. In spite of their poor characteristics within the transmitting band it has hitherto been necessary to use mid-shunt image impedance terminations of the "constant-k" type in these circuits. Impedance correcting sections of the first type are not suitable for this service because the complicated susceptance and reactance annulling networks at their line terminals produce sharp changes in reactance in the attenuating region. The outermost branch of terminations of the second type, however, is of simple configuration and if we choose it to resemble the final branch of a mid-shunt terminated "constant-k" type filter, as has been done in the sections shown in Figs. 16 and 17, we will secure an impedance characteristic beyond the band almost as good as that of the "constant-k" filter. Within the transmitting band, of course, its impedance is much better than that of the normal filter section.

Attenuation Characteristics of Terminating Sections

In the practical application of either type of terminating section some others of their characteristics, such as their transmitting efficiency and the effect produced upon them by parasitic dissipation of energy in the network elements, are also of importance. The transmission characteristics of the networks can be determined roughly by comparing them with standard filter sections. Let us consider, for example, the two branch termination shown in Fig. 20. If we neglect for

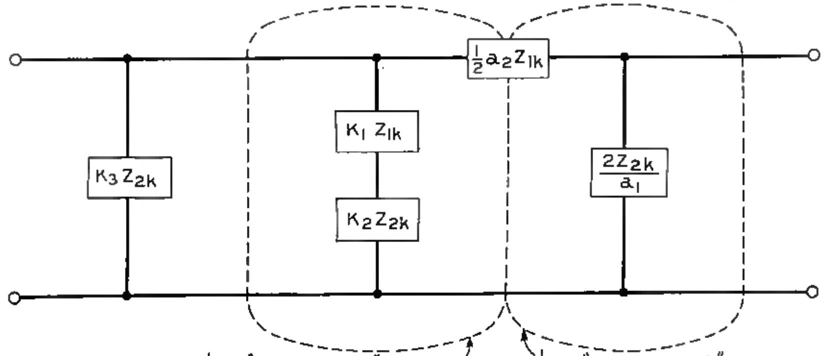


Fig. 20—Figure illustrating approximate transmission characteristics of 2-branch terminations.

the moment the third element of the susceptance correcting network, the remainder of the structure can be divided, in the manner indicated by the broken lines, into two portions, one of which resembles half of a "constant-k" section and the other half of an "m-derived" section in physical configuration. The transmission characteristic of the actual network is substantially similar to that which would be furnished by standard filter sections of these types. The mere fact that the network functions as an impedance corrector is, of course, sufficient to show that it will transmit efficiently frequencies within the nominal transmission band of the filter. Beyond the transmission band the attenuation characteristic would be almost exactly coincident with that of the suggested filter equivalent if it were not for the extra element in the final shunt branch. The extra element produces an anti-resonance in this arm somewhat beyond the resonance and near the anti-resonance point the attenuation is somewhat less than that which would be secured from ordinary filter sections. On the other hand the extra element considerably increases the admittance of the final shunt arm, and therefore the attenuation of the network, at frequencies remote from the cutoff. In spite of these modifications the analogy to standard sections is a fairly trustworthy guide to the attenuation of the networks. Several examples are given in the accompanying paper.

Since the ideal pure reactances contemplated by the theory are not physically available these conclusions must be modified somewhat in practical designs. As we might expect, however, unavoidable dissipation of energy in the network elements will alter the transmission characteristic of the correcting device about as it would that of an ordinary filter. In the attenuating range the effect can be neglected. In the nominal transmission band absorption of energy in the termination will reduce the transmitting efficiency of the circuit somewhat, but the loss in efficiency is no more serious than it would be in standard filter sections having the same general configuration.

Parasitic resistances in the network elements may of course affect the impedance as well as the transmission properties of the circuit. Since the structure is used primarily because of the impedance characteristic it furnishes, possible changes in impedance, caused by variations in the phase angles of the network elements, are of particular interest. Changes in impedance produced by dissipation of energy in the correcting networks, are easily estimated when the complete circuit with whose impedance we are concerned can be considered as a network of ordinary resistances, inductances and capacities and when dissipation affects the phase angles of all reactive elements equally. It can be shown that in such a network the change produced by dissipation in the resistance of the structure is proportional, to a first approximation, to the derivative of its reactance characteristic with respect to frequency, and that conversely the change in the reactance characteristic is proportional to the frequency derivative of the resistance characteristic. The explicit formulæ are:

$$\Delta R = \quad fd\frac{dX}{df},$$

$$\Delta X = -fd\frac{dR}{df},$$

where f is frequency and d the dissipation constant (defined as ratio of resistance to reactance) for each reactive element.

A filter, with its terminating sections and load resistance, is a network of resistances inductances and capacities to which the theorem applies. It seldom happens of course, that all of the reactive elements of the structure actually have the same dissipation constant. It is usually sufficient, however, to assume that "d" in the above formulæ is the average of the dissipation constants for coils and condensers. When well designed impedance correcting networks are used the reactance and resistance characteristics of the structure will be approximately constant over the operating range. The derivatives occurring in the above formulæ will consequently reflect only the presence of slight ripples in these characteristics about their mean values. The slopes of these ripples will usually be quite small. We can therefore conclude that *moderate amounts of dissipation will have no appreciable effect upon the impedance of a properly terminated filter*. The chief exceptions to this rule occur in low pass filters, where, at low frequencies the assumption that the dissipation constant is small is no longer satisfied.

In attempting to extend this principle to broader problems in impedance correction it is, of course, necessary to bear in mind that the analysis holds only for networks of resistances, inductances and capacities. We cannot expect the same results when the load impedance of the circuit has some arbitrary variation with frequency. For example, if we take the load impedance as the image impedance of a dissipationless "constant-k" filter and assume that parasitic resistances occur only in the termination, we will find that dissipation does change the impedance of the circuit. The circuit impedance will be insensitive to dissipation only when we include the complete structure, and not merely the terminations, in our analysis.

More General Problems of Impedance Correction

This general method of impedance correction having worked with reasonable success in its application to "constant-k" wave filter impedances, it is natural to inquire whether it can be applied to other problems with equal ease. Further possibilities for example might include the correction of other types of filter impedances, or the correction over extremely wide frequency bands for the effects of leakage inductance and finite mutual inductance in transformers, or the reduction of actual transmission line impedances to constant resistances. All of these possible applications assume that the impedance correcting device is a 4-terminal network, transmitting useful signal energy to its load impedance. When terminated by such an element as a simple resistance, however, it might also be used as a 2-terminal network, forming one branch of a complete system. By appropriate adjustment of the impedance controlling parameters the network could, theoretically at least, be given a wide range of impedance characteristics.

We might, for example, use it to approximate a pure resistance varying in an arbitrary manner with frequency, which would be a valuable impedance element in certain circumstances.

None of these possibilities has been investigated in detail, and naturally the measure of success which can be achieved with any one of them will depend largely upon the precise conditions of the problem. The mathematical form we have specified for the load impedance of the network is so broad however that if we were to consider only this aspect of the situation we might conclude that the scope of the structure is well nigh universal. For example, the impedance of any finite network of resistances, inductances, and capacities can be written in the appropriate mathematical form. Even when the load impedance is not described in the required manner, either because it is empirically determined or because it has the wrong theoretical formula, the type of algebraic expression we have been considering is so general that it can always be matched approximately.

Unfortunately, the range of application promised by this rather superficial mathematical discussion may be severely restricted by other considerations. In the general case, for instance, the number of terms in the denominator of the resistance expression will be greater than the number of branches in the correcting network and it will not be possible to choose them all arbitrarily. Moreover, even when the correct number of conditions have been imposed upon the power series coefficients we have no assurance that the resulting system of simultaneous equations between coefficients and element values can be solved, or that the solutions, if obtained, will always correspond to physically realizable elements. Finally, we may observe that although no difficulty was experienced in the reactance or susceptance correction of filters, it seems probable that, in view of the limited range of characteristics which can be simulated by physically realizable reactive structures, a straightforward application of the general method of resistance correction will often leave us with a reactive characteristic which cannot be corrected.

These difficulties may occasionally be overcome by slight modifications in the design process. Among other possibilities for example, we can adjust the lowest powers both in the denominator of the resistance expression and numerator of the reactance expression [9] to desirable values, obtain an approximate value for the effect of higher powers in both expressions by a trial computation and readjust the coefficients of the lower powers to take account of these previously neglected terms.

[9] Since the denominator of the reactance is equal to that of the resistance, whose value is prescribed by the requirements, the reactance expression can be determined completely from its numerator alone.

Difficulties appearing in a direct application of the impedance correcting process may also be avoided if we adopt the reverse method of impedance correction suggested by the theorem on reciprocal impedance relationships. The method has already been applied to the construction of alternative filter impedance correcting sections. Similar alternative configurations can be built up in any impedance correcting problem if we consider that the structure is terminated by the conjugate of the desired impedance and adjust its parameters to produce the conjugate of the given impedance. Since the desired impedance will in general be a relatively simple function of frequency, this alternative procedure at least avoids analytical complexities. In spite of these possibilities however it seems probable that the method will fail in many situations. It seems best adapted to such problems as that of filter impedance correction, where a transformation must be made from one fairly simple characteristic to another simple characteristic. An attempt to apply it to more difficult problems should result, at best, in very complicated networks.

Transmission Properties of Impedance Correcting Networks

The close relationship between the impedance correcting properties of our networks and their transmission characteristics has been manifest from time to time in the previous discussion. The networks used at filter terminations, for example, transmitted freely within the range in which they functioned satisfactorily as impedance correctors but attenuated other frequencies. That this will be true in general is easily seen by inspection. Within the range in which a desired impedance characteristic is obtained, of course, our previous argument from the principle of conservation of energy is alone sufficient to show that the networks transmit with the maximum efficiency compatible with the impedance requirements imposed upon the circuit. On the other hand, it is evident from the filter-like configuration of the networks that at frequencies remote from the operating range of the networks, where the parameter "x" becomes large, the structures will ordinarily introduce attenuation. From the impedance standpoint this means merely that for sufficiently large values of x the polynomial approximations upon which the analysis is based no longer hold, and the resulting mismatch between the generator and network impedances diminishes the amount of power which can enter the structure.

When the impedance correcting analysis is stated in a slightly modified form, whose possibilities have not as yet been completely investigated, the impedance and transmission characteristics of the circuit are still more firmly related. Thus for example the attenuation of the structure beyond its operating range results chiefly from the readily computed departure of the resistance or conductance characteristic of the network from that of the generator. It is also produced, in part, however, by the failure of the reactance or susceptance correcting network to annul in this range the imaginary component furnished by the resistance or inductance controlling network and the effect of this factor is less easy to determine. In the modified analysis it is often possible to do away with the distinction between the two types of networks. The complete insertion loss characteristic is then embodied in a single polynomial expression. In the modified form, moreover, the analysis may often be used to determine the phase as well as the attenuation of the circuit.

Granted these results, it is but a short step to the conclusion that the impedance correcting analysis offers a possible approach to the design of filters. While it is usually true that the networks will attenuate frequencies beyond the region in which impedance requirements have been set, the amount of the mismatch which produces this attenuation, since it depends upon the impedance correcting parameters, is still more or less under our control. By suitable adjustments of the correcting network, therefore, we can design a structure to meet attenuation as well as impedance requirements. A particularly interesting situation occurs when the load impedance is a constant pure resistance.[10] As we have already seen, a load impedance of this type satisfies our mathematical specification and it can therefore be used with a ladder network. Since a perfect impedance match already exists in the circuit an inserted network can be called an impedance correcting device only by courtesy. Unless the network contains so many branches that the mathematical complexity of the problem is overwhelming, however, it is possible to so manipulate the impedance correcting parameters that the network impedance matches the generator impedance approximately over a certain frequency band but is very poor outside this range. It follows from our previous discussion that the network will transmit frequencies lying within this band efficiently, but will attenuate other frequencies. Networks designed in accordance with this method therefore function as filters. They differ from conventional filters in several respects, however. For example they are non-recurrent, they cannot be divided into discrete sections with matched image impedances, and they do not possess definite cutoffs.

[10] This circuit arrangement was first investigated by E. L. Norton and W. R. Bennett, who developed a complete analysis for a number of particular cases.

Abstracts of Technical Articles from Bell System Sources

A Space-Time Pattern Theory of Hearing.[1] HARVEY FLETCHER. The pitch of a tone is determined both by the position of its maximum stimulation on the basilar membrane and also by the time pattern sent to the brain. The former is probably more important for the high tones and the latter for the low tones. The loudness is dependent upon the number of nerve impulses per second reaching the brain and possibly somewhat upon the extent of the stimulated patch. The experience called by psychologists "volume" or "extension" is no doubt identified with the length of the stimulated patch on the basilar membrane. This extension is carried to the brain and forms a portion of excited brain matter of a definite size. It is then this size that determines our sensation of the "volume" of a tone. The low pitched or complex tones have a large "volume" while the high pitched tones have a small one.

The psychological experience called "brightness" may be identified with the sharpness of the peaks in the vibration form of the basilar membrane as suggested by Dr. Troland. The high tones give the sense of brightness while the low tones the sense of dullness.

The time pattern in the air is converted into a space pattern on the basilar membrane. The nerve endings are excited in such a way that this space pattern is transferred to the brain and produces two similar space patterns in the brain, one on the left and the other on the right side. Enough of the time pattern in the air is sent to each of these stimulated patches to make times of maximum stimulation in each patch detectable. So when listening to a sound with both ears, there are four space patterns in the brain produced, each carrying also some sort of time pattern. It is a recognition of the changes in these patterns that accounts for all the phenomena of audition.

The Theory of Probability: Some Comments on Laplace's Théorie Analytique.[2] E. C. MOLINA. This paper is concerned with an answer to the questions "to what extent will one conversant with the *Théorie Analytique* be in touch with the present status of probability theory, and how sound a foundation will he have found therein for statistical applications of the theory?"

[1] *Jour. Acous. Soc. Amer.*, April, 1930.
[2] *Bulletin, Amer. Math. Soc.*, June, 1930.

In answer to the first question emphasis is laid on the virtual identity between Laplace's *generating function* and the Cauchy-Poincaire *characteristic function*, on the close approach of Laplace's analysis to the form of the Fourier *reciprocal* equations and to the explicit presentation by Laplace of the *Hermite polynomials* and related *Gram-Charlier* expansion. In answer to the second question, the author submits Laplace's contributions to the probability of causes and points out the distinction drawn by Laplace between the meaning of the word *limit* when used outside the domain of probability theory and its meaning when the word is attached to the observed frequency with which an event happens.

As evidence that the *Théorie Analytique* is in advance of much recent literature, and on account of its great practical value, the Laplacian method of dealing with integrands involving factors raised to high powers is outlined. In this connection attention is called to a Laplacian differential equation which contains, as a special case, the differential equation from which Karl Pearson has derived his famous system of frequency curves.

Method of Enhancing the Sensitiveness of Alkali Metal Photoelectric Cells.[3] A. R. OLPIN. A technique is described for sensitizing alkali metal photoelectric cells to light by introducing onto the metal surface small amounts of dielectrics, as oxygen, water vapor, sulphur vapor, sulphur dioxide, hydrogen sulphide, air, sodium bisulphite, carbon bisulphide, etc., or some organic compound as methyl alcohol, acetic acid, benzene, nitrobenzene, acetone, etc., or some organic dye as tropæolin, rosaniline base, eosin, cyanine, kryptocyanine, dicyanine, neocyanine, etc. The marked increase in electron emission from the cathodes of cells so treated is due primarily to an increase in response to red and infrared light. Vacuum sodium cells have been produced, yielding photoelectric currents as high as 7 microamperes per lumen of white light of color temperature 2848° K and cæsium cells yielding far greater currents.

The response of these cells is proportional to the intensity of the exciting light even for light of longer wave-lengths than that to which the cell responded before treatment.

Spectral response curves are similar for all cells using the same metal as cathode. These curves differ from the curves for the pure metal by the appearance of a new selective maximum at lower frequencies. This newly appearing maximum resembles the regular maximum for the untreated metal and is due to the presence of the sulphur and air. Changes of approximately 0.8 volt are common.

[3] *Phys. Rev.*, July 15, 1930.

The validity of Einstein's equation precludes the possibility of explaining the new maximum in the spectral response curve for a treated surface by a "Raman shift" of the incident light frequencies, even though the separation of these maxima is equal to certain well-known vibration-rotation frequencies of the dielectric molecules. It may be that the natural frequency of the alkali metal atom is diminished by the vibration frequency of the complex atom in which it is incorporated.

The Lindemann formula for the frequency of the selective photoelectric maximum $[2\pi\nu = (ne^2/mr^3)^{\frac{1}{2}}]$, primitive though it seems in the light of modern theory, has always given values for the pure metals in close agreement with experimental determinations. The n term is determined by the valence of the substance, a choice of unity being used for the monovalent alkali metals corresponding to an electron revolving around a singly charged ion. A choice of 2, 3,—for divalent, trivalent,—substances corresponds to electrons revolving around doubly, triply-charged ions. Under certain conditions the alkali metals manifest different valencies, such for instance, as those exhibited in the oxide series Na_2O_2, Na_2O, Na_3O, Na_4O. These compounds can be prepared in vacuum and are light-sensitive. Spectral response curves for such cells exhibit all the selective maxima always separated from it by the frequency of a well-known line in the vibration-rotation spectrum of the dielectric molecules, usually the 1.5μ line so characteristic of oxygen-hydrogen, carbon-hydrogen or nitrogen-hydrogen linkages. The long wave limit shifts an amount agreeing with the separation of the maxima.

With a cell so designed that the cathode could be sensitized in a side chamber and then slipped into its proper place (thus keeping the anode free from light-sensitive materials), stopping potentials were obtained for electrons, liberated by monochromatic light, from a sodium cathode before and after treating it with sulphur vapor and air. For light of wave-lengths ranging from λ3500A to λ8000A falling on the treated cathode, the electron retarding potentials are found to vary linearly with the frequency of the exciting light, thus establishing the validity of Einstein's photoelectric equation for composite surfaces. From the slope of the straight line depicting this relationship, the value of Planck's constant h is found to be 6.541×10^{-27}, significant to three figures. An almost identical value is obtained for untreated sodium. The apparent stopping potentials, or voltages at which the photoelectric currents become zero are the same before and after the sulphur and air treatment. The voltage at which the current just saturates is always greater after treatment than before. This is a measure of the change in contact potential of the cathode called for by the Lindemann formula when the value of n is chosen to agree with the valence of the metal. Data are presented showing this condition to be general for the alkali metals, a maximum response to red or infrared light being dependent upon the formation of a subvalent compound, as a suboxide.

Attention is called to seemingly analogous phenomena in the fields of photoelectricity, photography, fluorescence and absorption.

Some Problems in Short-Wave Telephone Transmission.[4] J. C. SCHELLENG. In this paper are discussed certain phases of short-wave telephony, primarily, though not entirely, from the point of view of the transmitter. The field strengths which the transmitting station must provide at the receiver are considered. Typical data are given showing results obtained in transmission from Deal, New Jersey, to England. This is followed by a discussion of requirements and limitations of the transmitting antenna. The gains which arrays may reasonably be expected to provide are considered. The phenomenon of non-synchronous fading at nearby points is examined as to its bearing on the dimensions and performance of directive arrays. Other directional properties of the transmitting medium are also considered. Attention is then directed to the transmitting equipment, particular attention being given to the high-power part of it. Requirements, rather than circuit details, are emphasized. These include stability of operation, flexibility, and freedom from amplitude distortion, and phase and frequency modulation. The results of tests in which some of these matters were considered quantitatively are given.

A Chronographic Method of Measuring Reverberation Time.[5] E. C. WENTE and E. H. BEDELL. Reverberation time measurements are generally made with the ear and a stop watch in the manner devised by Prof. Wallace Sabine. Surprisingly consistent results can be obtained by this method in a reverberation chamber, where the rate of decay of sound is slow and where disturbing sounds are absent. But such measurements present difficulties if the room is noisy or if the reverberation time is short. Also it is recognized that uncertainties may be introduced because of the fact that the threshold of hearing varies between individuals and with time in the same person. It was with the object of overcoming these difficulties that the electrical method described in this paper was devised. This method does not differ essentially from that of Sabine except that an electro-acoustical ear of controllable threshold sensibility is substituted for the human ear.

[4] *Proc. I.R.E.*, June, 1930.
[5] *Jour. Acous. Soc. Amer.*, April, 1930.

Contributors to this Issue

H. W. Bode, A.B., Ohio State University, 1924; M.A., 1926. Bell Telephone Laboratories, 1926–. Mr. Bode has been engaged in the study of transmission networks, such as wave filters, attenuation equalizers, and phase correctors.

A. E. Bowen, Ph.B., Yale University, 1921; Instructor, Department of Physics, Yale University, 1921–24; American Telephone and Telegraph Company, Department of Development and Research, 1924–. Mr. Bowen has been engaged mainly in work on problems of inductive coordination of power and communication systems.

Karl K. Darrow, B.S., University of Chicago, 1911; University of Paris, 1911–12; University of Berlin, 1912; Ph.D., University of Chicago, 1917; Western Electric Company, 1917–25; Bell Telephone Laboratories, 1925–. Dr. Darrow has been engaged largely in writing on various fields of physics and the allied sciences. Some of his earlier articles on Contemporary Physics form the nucleus of a recently published book entitled "Introduction to Contemporary Physics" (D. Van Nostrand Company). A recent article has been translated and published in Germany under the title "Einleitung in die Wellenmechanik."

Ward F. Davidson, B.S.E., University of Michigan, 1913; M.Sc., 1920; Westinghouse Electric and Manufacturing Company, 1914–16; University of Michigan, teaching electrical engineering, 1916–22 except for two years war duty; Brooklyn Edison Company, Inc., Director of Research, 1922–. He has devoted much time to theoretical and experimental study of underground power transmission cables.

C. L. Gilkeson, B.S., Massachusetts Institute of Technology, 1922, M.S., 1923; Transmission and Distribution Department of the Philadelphia Electric Company, 1923–26; National Electric Light Association 1926–. Mr. Gilkeson's work with the N.E.L.A. has been in connection with the research program of the Joint Subcommittee on Development and Research of the N.E.L.A. and Bell System. Much of his work has been associated with problems of low frequency induction.

E. I. Green, A.B., Westminster College (Fulton, Mo.), 1915; University of Chicago, 1915–16; U. S. Army, 1917–19 (Captain, Infantry); B.S. in Electrical Engineering, Harvard University, 1921; Department of Development and Research, American Telephone and Telegraph Company, 1921–. Mr. Green has been engaged principally in work on line transmission problems and multiplex transmission systems.

R. G. McCurdy, B.S., University of California, 1913; Technical staff, Joint Committee on Inductive Interference of the Railroad Commission of California, 1913–16; American Telephone and Telegraph Company, Engineering Department, 1916–19; Department of Development and Research, 1919–; Noise Prevention Engineer, 1930. Mr. McCurdy's work has been chiefly on problems of inductive coordination of telephone and power circuits.

W. H. Martin, A.B., Johns Hopkins University, 1909; B.Sc., Massachusetts Institute of Technology, 1911; American Telephone and Telegraph Company, Engineering Department, 1911–19; Department of Development and Research, 1919–. Mr. Martin's work has related particularly to transmission of telephone sets and local exchange circuits, transmission quality and loading.

E. B. Payne, B.S. in E.E., Mass. Inst. of Technology, 1917; Artillery Corps, U. S. Army, 1917–18; Engineering Department, Western Electric Company, 1919–25; Bell Telephone Laboratories, 1925–. Mr. Payne has been engaged in loading coil design and more recently concerned with the development of wave filters and allied transmission networks.

Liss C. Peterson, E.E., Chalmers Technical Institute, Gothenburg, 1920; Technische Hochschule, Charlottenburg, 1920–21; Technische Hochschule, Dresden, 1921–22; Signal Corps, Swedish Army, 1922–23; American Telephone and Telegraph Company, 1925–30; Bell Telephone Laboratories 1930–. Mr. Peterson's work has been concerned mostly with problems relating to inductive interference.

R. R. Williams, B.S., University of Chicago, 1907, M.S., 1908; Research Chemist, Bureau of Science, Philippine Islands, 1908–15; Bureau of Chemistry, U. S. Department of Agriculture, 1915–18; Engineering Department, Western Electric Company, 1918–25; Bell Telephone Laboratories, 1925–. Mr. Williams has done extensive research work on submarine cable insulation. Since 1925, as Chemical Director, he has been in charge of the Chemical Laboratories of the Research Department of the Bell Telephone Laboratories.

W. J. Williams, C.E., Rensselaer Polytechnic Institute, 1905; Assistant Engineer with Dr. W. L. Robb, 1905–06; Instructor in Physics and Electrical Engineering, R.P.I., 1906–15; Professor of Electrical Engineering, R.P.I., 1915–. Mr. Williams is a Consulting Electrical Engineer and has been Technical Adviser for the National Electric Light Association since 1923, on work of the Joint Subcommittee on Development and Research of the N. E. L. A. and Bell System.

Leon T. Wilson, Ph.B., Yale, 1915; Graduate Student and Instructor, Yale, 1915–17 and 1920–22; Signal Corps, U. S. Army, 1917–19; E.E., Yale, 1919; Department of Development and Research, American Telephone and Telegraph Company, 1923–. Mr. Wilson's earlier work was chiefly in the radio field and included the development of a thermocouple type of voltmeter for radio frequencies. With the Department of Development and Research his work has dealt mainly with insulator problems of open-wire lines.