HEINRICH RUDOLPH HERTZ

1857–1894

The Bell System Technical Journal

Vol. XVII *July, 1938* *No. 3*

Hertz, the Discoverer of Electric Waves *

By **JULIAN BLANCHARD**

FIFTY years have passed since those memorable researches of the young German physicist, Heinrich Rudolph Hertz, which have come to be regarded as the starting point of radio. For it was he who first detected, and measured, electromagnetic waves in space—waves which had been predicted, it is true, but which had never before been observed. It is not to be claimed, of course, that the radio art would have failed to be born were it not for his genius, for we know that almost simultaneously the experiments of Lodge in England were pointing with certainty to the same discoveries, and the speculations of others were revolving around the possibility of generating electric waves. Yet it was the remarkably clear vision of Hertz, combined with his consummate persistence and skill, that won for him the prize and justly enshrined his name among the immortal men of science.

So, upon this golden anniversary of the opening of a new epoch in the realm of communication, it is fitting that we pause to do honor to his memory and to consider anew the significance of his great accomplishment.

The formal facts of Hertz's biography can be set down very quickly. He was born at Hamburg on February 22, 1857, his father an attorney, belonging to a family of successful merchants, his mother the daughter of a doctor of medicine, and the descendant of a long line of Lutheran ministers—all of cultural tastes and attainments on both sides. At the age of twenty he went off to school at Munich, after a rather unorthodox preparatory training, supposedly to pursue an engineering career, but he was torn between this resolve and his natural inclination for the study of pure science. Soon after reaching Munich he felt compelled to put the matter before his parents, to whom he frequently and confidingly wrote concerning his plans and his work. In a long letter written in November, 1877, he said, "I really feel ashamed to say it, but I must: now at the last moment I want to change all my

---

* Published in *Proc. I.R.E.*, May, 1938, Vol. 26, No. 5.

plans and return to the study of natural science. I feel that the time has come for me to decide either to devote myself to this entirely or else to say good-bye to it; for if I give up too much time to science in the future it will end in neglecting my professional studies and becoming a second-rate engineer. Only recently, in arranging my plan of studies, have I clearly seen this—so clearly that I can no longer feel any doubt about it. . . ." And then follows a lofty and appealing presentation of the reasons for his choice. Concluding, he wrote: "And so I ask you, dear father, for your decision rather than for your advice; for it isn't advice that I need, and there is scarcely time for it now. If you will allow me to study natural science I shall take it as a great kindness on your part, and whatever diligence and love can do in the matter, that they shall do. I believe this will be your decision, for you have never put a stone in my path, and I think you have often looked with pleasure on my scientific studies . . ."

Matters were arranged as he wished and he joyfully pursued his studies at the University and at the Technical Institute, working hard on mathematics and mechanics, and spending much time in the physical laboratory. In October, 1878, he went to Berlin and became there a student under the mighty giants, von Helmholtz and Kirchhoff; writing to his parents in November, "I am now thoroughly happy, and could not wish things better." In 1880 he gained his doctorate, and in October of that year was appointed assistant to Helmholtz, which delightful and stimulating association continued until Easter of 1883. He then went to the University of Kiel to become lecturer in theoretical physics, and here he first began to reflect seriously upon Maxwell's electromagnetic theory of light.

He was soon promoted again, and at Easter, 1885, he became professor of experimental physics at the Technical High School in Karlsruhe. Here, in 1886, at the age of twenty-nine, he married Elizabeth Doll, daughter of the professor of geodesy in the same institution, and their home became a congenial meeting place for their many cultured associates. It was at Karlsruhe also that he began his researches on electric waves. Before they were finished he was called in 1889 to succeed the celebrated Clausius at the University of Bonn, thus at the age of thirty-two having arrived at a position in the academic world not ordinarily to be achieved until much later in life.

He soon thereafter relinquished to others the further exploration of the great new territory of electric waves he had opened up and returned to some investigations on the discharge of electricity in rarefied gases, a subject which had interested him while at Berlin. He then devoted his attention entirely to what proved to be his last work, a treatise on mechanics. In the summer of 1892 he suffered a severe illness which eventually led to chronic blood poisoning, of which he died, after indescribable suffering, on January 1, 1894. He would have been just thirty-seven years old the following month.

Although the fame of Hertz rests primarily on his electric wave researches, these constitute by no means the whole of his work. His collected papers, edited by the German physicist, Dr. Philipp Lenard, and admirably translated into English by Professor D. E. Jones and associates, comprise three volumes. The first consists mainly of his miscellaneous earlier papers, some twenty-odd titles altogether. One of these, published in 1884, "On the Relations between Maxwell's Fundamental Electromagnetic Equations and the Fundamental Equations of the Opposing Electromagnetics," marked an important step in the development of Hertz's ideas, and has been called his greatest contribution to theoretical physics. In it he opposed the old orthodox theories of electric phenomena based on action at a distance, which were supported by most of the Continental physicists, and definitely aligned himself with the followers of Maxwell. This volume also contains his semipopular Heidelberg lecture of 1889, "On the Relations between Light and Electricity," giving a general account of his more recent work. It strikingly illustrates the charm and felicity of style he could employ in the presentation of a difficult subject, and cannot fail to be read, and reread, with pleasure and admiration. The volume ends with a eulogy of Helmholtz on the occasion of his seventieth birthday, wherein the pupil, in equally graceful language, paid homage to his beloved and inspiring preceptor. These two papers reveal so clearly between the lines the manner of man their author was.

The second volume contains the papers on electric waves, which had been collected by the author himself, as a result of numerous requests for reprints, and published under the German title, "Untersuchungen über die Ausbreitung der Elektrischen Kraft," and later in English as "Electric Waves," with a lucid introduction by Hertz explaining the motive and significance of each of the separate papers. One of the first of these describes an important by-product of the electric wave studies, the discovery of the effect of ultraviolet light upon an electrical discharge. This discovery itself uncovered a new wealth of physical problems and the subject became immediately of great interest to many experimenters, most of whom at the time little suspected that this subsidiary effect was not the main discovery, or imagined that electrical science was on the eve of much greater conquests. The paper attracted attention to Hertz and aroused a popular interest in him, so that everything coming from him thereafter was eagerly read. Had it not been for this rather accidental circumstance the importance of his subsequent work might have gone temporarily unnoticed, as has happened with some of the greatest discoveries.

The third volume in the series is his book, "Die Prinzipien der Mechanik," completed with difficulty during his last illness and published a few months after his death. A lengthy preface by the venerable Helmholtz gives an appreciative sketch of the author's life and work—this time a sorrowful tribute from master to pupil.

In order to understand and appraise the work of Hertz on electric waves, it will be needful to review briefly the ideas about light and electricity that prevailed at his time and before. With regard to light, Newton's corpuscular theory had given way in the early 1800's to the wave theory of Young and Fresnel, and long before Hertz's day the idea of transverse vibrations in a hypothetical elastic-solid medium called the luminiferous ether had become firmly established. Length of wave and velocity had been measured many times. The wave theory accounted satisfactorily for all the known phenomena in optics and there was no doubt in anybody's mind about its essential correctness, regardless of any difficulties encountered in explaining the nature of the ether and its relation to matter.

As to electricity and magnetism, the older theories of instantaneous action at a distance were beginning to weaken with the discoveries, in the early part of the nineteenth century, of the reactions between electric currents and magnets, and the phenomenon of induction. Hitherto there had been no postulation of an intervening medium to explain the transmission of the force between two charged bodies, and it was supposed that electricity and magnetism, like gravitation, acted across empty space in straight lines and instantaneously. In somewhat different dress these ideas were given new life around the middle of the century, particularly by some of the German physicists. To Faraday, however, such views were unacceptable. He wished to get rid of the idea of action at a distance and in his mind pictured a medium, along the contiguous molecules or particles of which the force was propagated. In this medium he visualized "lines of force" emanating from or terminating upon the electric charges or magnetic poles, acting like stretched elastic threads, repelling each other sidewise as well as tending to contract. Thus, in his thinking, attention was focused upon the insulating medium surrounding a conductor, the "dielectric" as he called it, for here, he thought, was the real seat of the action. He believed also that there existed some direct relation between light and electricity or magnetism. He was ever seeking to find such a relation and in the course of his many experiments he discovered the rotation of the plane of polarization of light by a magnetic field. In his speculations there was one question which was continually presenting itself to him. Do electric and magnetic forces, like light, require time for their propagation? Are there waves? But to this question he was unable to find an answer.

And then came Maxwell, building upon the foundation that Faraday had laid, translating Faraday's ideas into the language of mathematics, and making the grand generalization that light and electric waves are one and the same phenomenon, propagated by the same medium, with the same velocity, and differing only in wave-length. Like Faraday, he considered the energy of the electromagnetic field to reside in the dielectric. He conceived the medium to have properties analogous to those of an elastic solid, which would spring back to its original state upon the removal of the straining force. The alteration of the displacement, or "polarization," in the medium was viewed by him as an electric current, which he called the "displacement current," as distinguished from the "conduction current" existing in conductors. From the general equations which he formulated it was shown that only transverse vibrations (like light) could be propagated in such a medium and that the velocity of propagation should be equal to the ratio of the two systems of electrical units, that is, to the number of electrostatic units of electricity contained in one electromagnetic unit. This ratio had been experimentally determined by Weber and Kohlrausch (it was later redetermined by Maxwell himself, by a different method), and the fact that it agreed so very closely with the measured velocity of light was one of the strongest points in favor of the view that light waves were identical with the hypothetical electromagnetic waves. Maxwell's comprehensive theory was first enunciated in his 1865 paper entitled "A Dynamical Theory of the Electromagnetic Field," and afterwards elaborated in his great "Treatise on Electricity and Magnetism," published in 1873, but for a number of years it was regarded by many as merely a speculation, by others as probably true, and by none as conclusively proved. It remained for Hertz to add the capstone to the theory by actually demonstrating for the first time the existence of electromagnetic waves in space; and furthermore, to show experimentally that they had all the physical properties of ordinary light waves.

In 1888, while he was teaching at the Technical High School in Karlsruhe, Hertz carried out the brilliant experiments which have made his name famous. These were actually a part of a long series of experiments which began in 1886, and which came about partly by accident, and yet were the result of his keen interest in everything connected with electric oscillations; an interest extending back to 1879, when, at the suggestion of Helmholtz, he had considered tackling a prize problem proposed by the Berlin Academy of Science aimed at the proof of a portion of Maxwell's theory, but which he had abandoned for the reason that oscillations of sufficiently high frequency were not then available. While using in his lectures at Karlsruhe a pair of flat ("pancake") coils, called Reiss or Knockenhauer spirals, mounted adjacent to each other, he was surprised to find how easy it was to obtain sparks between the terminals of the secondary coil when a small Leyden jar or even a small induction coil was discharged through the primary, provided the primary discharge took place across a spark gap. This, of course, was an indication of an exceptionally strong inductive effect. This observation led to his discovery of a method of exciting electric disturbances far more rapid than any hitherto known, such as those of Leyden jars or open induction coils as customarily used, and having wave-lengths, it turned out, capable of being measured within the confines of a laboratory. His oscillator was nothing more than a short metal rod with a spark gap in the middle (sometimes with metal spheres or plates attached to the ends, resembling a dumbbell), the sparking terminals consisting of small knobs or spheres which were connected to the terminals of a Ruhmkorff induction coil; the small inductance and capacitance of this simple linear conductor, together with the proper functioning of the spark gap, accounting for its success. By such means Hertz obtained wave-lengths from a few meters down to 30 centimeters, and so began, it is seen, with the "ultra-short" waves that are again coming into vogue.

Hertz began his experiments with a study of the "induction" about this exciter. As he commented in his first paper in this series, "On Very Rapid Oscillations," published in May, 1887, theory had predicted the possibility of very rapid oscillations in open-wire circuits of small capacitance, but it could not be predicted from theory whether they could be produced on such a scale as to admit of their being observed. Hertz not only devised a method of producing such oscillations, but also discovered a method of detecting them, by their effects in the surrounding space. His detector consisted merely of a short length of wire bent in the form of a rectangle or a circle and containing a micrometer spark gap, across which minute sparks could be seen in a darkened room; especially if this secondary circuit was in electrical resonance with the exciter. This exceedingly simple detector was indeed a capital discovery. Some five years earlier Professor G. F. Fitzgerald, of Dublin, had suggested "the combination of a vibrating generating circuit with a resonant receiving circuit . . . as one by which this very question might be studied." But, as he said afterwards in speaking of Hertz's work, "I did not see any feasible way of detecting the induced resonance: I did not anticipate that it could produce sparks." Concerning this contrivance the following interesting remarks were made by its author in his Heidelberg lecture above referred to: "The method had to be found by experience, for no amount of thought could well have enabled one to predict that it would work satisfactorily. For the sparks are microscopically short, scarcely a hundredth of a millimeter long; they only last about a millionth of a second. It seems absurd and almost impossible that they should be visible; but in a perfectly darkened room they *are* visible to an eye which has been rested in the dark. Upon this thin thread hangs the success of our undertaking." *Multum in parvo*, truly!

After a series of preliminary experiments, in which he studied the various induction effects, including the phenomenon of resonance, and demonstrated waves on wires (an earlier, but overlooked, discovery of von Bezold, in 1870), and also solved the problem of the Berlin Academy—"to establish experimentally any relation between electromagnetic forces and the dielectric polarization of insulators"—he was fully convinced that the disturbance was propagated through space, independently of wires, with a finite velocity and in the form of waves, in accordance with Maxwell's prediction. His conclusion was then definitely and convincingly proved by making use of the well-known method of reflection and interference to produce standing waves, and noting the position of the nodes and antinodes. These epochal experiments were described in a paper entitled "Electromagnetic Waves in Air and Their Reflection," published in May, 1888. But he did not stop there, and in these and succeeding investigations he showed that electric waves are reflected from plane and curved metal surfaces in accordance with the same laws as light waves; that they are refracted in passing thorugh prisms of pitch, paraffin, and other dielectrics; and that they are polarized by a grating of parallel wires, and hence are transverse waves. From actual measurements of their wave-length and computations of their frequency (from the constants of his oscillator), he calculated their velocity and found that it was the same as the velocity of light. As summarized by Hertz himself, "The object of these experiments was to test the fundamental hypotheses of the Faraday-Maxwell theory, and the result of the experiments is to confirm the fundamental hypotheses of the theory." The old action-at-a-distance philosophy had come to an end.

The importance of Hertz's contributions to this great subject received instant and enthusiastic recognition, and his experiments were soon being repeated in all the important laboratories of the world. The English mathematical physicist, Oliver Heaviside, writing in 1891, said: "Three years ago electromagnetic waves were nowhere. Shortly after, they were everywhere." Here were researches of a most abstruse and complex character, with no apparent utility and having no elements of popular appeal, and yet bringing to their author such acclaim as had seldom been accorded to a man of science. Honors were showered upon him on every hand, at home and abroad. In England, where his work was especially appreciated, he was awarded the coveted Rumford medal by the Royal Society.

Hertz's characteristic modesty in referring to his own achievements was matched only by his generosity in giving credit to the accomplishments of others. In one of his lectures he said, "Such researches as I have made upon this subject form but a link in a long chain. . . . Lack of time compels me, against my will, to pass by the researches made by many other investigators; so that I am not able to show you in how many ways the path was prepared for my experiments, and how near several investigators came to performing these experiments themselves." Mention has been made of the investigations of Sir Oliver Lodge in the same field and the imminence of his discovery of the same phenomena. It is pleasant, indeed, in this instance to be able to record the absolute lack of any feeling of jealousy or envy on the part of either of these courteous gentlemen. In the introduction to his collected papers Hertz wrote, "I may here be permitted to record the good work done by two English colleagues who at the same time as myself were striving towards the same end. In the same year in which I carried out the above research, Professor Oliver Lodge, in Liverpool, investigated the theory of the lightning conductor, and in connection with this carried out a series of experiments on the discharge of small condensers which led him on to the observation of oscillations and waves in wires. Inasmuch as he entirely accepted Maxwell's views and eagerly strove to verify them, there can scarcely be any doubt that if I had not anticipated him he would have succeeded in observing waves in air, and thus also in proving the propagation with time of electric force. Professor Fitzgerald, in Dublin, had some years before endeavored to predict, with the aid of theory, the possibility of such waves, and to discover the conditions for producing them."

On his part Lodge just as generously wrote, only a few years afterwards, in an obituary of his rival: "Hertz stepped in before the English physicists, and brilliantly carried off the prize. He was naturally and unaffectedly pleased with the reception of his discovery in England, and his speech on the occasion of the bestowal of the Rumford medal by the Royal Society will long be remembered by those who heard it for its simplehearted enthusiasm and good-feeling. His letters are full of the same sentiment. . . ."

Noteworthy, indeed, was the extreme modesty of this scientific lion of the hour, and equally striking his consideration for the feelings of others. It is recorded that when the Royal Society presented him with the Rumford medal, "he silently disappeared from Bonn for a few days —none knew why—and he returned as silently." In refusing the request made by the editor of *The Electrician* (of London) in 1890 for his photograph, Hertz replied, "I feel as if presenting my portrait now in so prominent a place follows too quickly the little work I have done. I should like to wait a little, and see if the general approbation which my work meets with is of a lasting kind. Too much honor certainly does me harm in the eyes of reasonable men, as I have sometimes occasion to observe. If your kind intention is the same in two years, even one year, I shall readily consent and help you in every respect." Four years later the portrait was published, following upon Hertz's death.

Upon the untimely ending of his brief but brilliant career, occurring in the very prime of life, before he was yet thirty-seven, there was a feeling of shock and sadness in every scientific quarter. Many were the sincere tributes paid to his memory, honoring him for his rare personal qualities as well as his distinguished scientific attainments. Some expressions from Lodge have already been quoted. Said he in his obituary in *The Electrician*, "Not a student of physical science on the planet but will realize and lament the sad loss conveyed by the message, 'Hertz is dead.'" The editor of that journal wrote, "In the modesty and self-forgetfulness which blend so admirably in the spirit of true scientific research Hertz was singularly rich." In an editorial note in an American journal, *The Physical Review*, we find the following: "In addition to the recognition of those who were able to appreciate his work, Hertz received the acclamations of the entire world of thought. Fortunately he possessed a nature of such complete simple-mindedness that his sudden rise into a position akin to notoriety had no effect upon him. The unassuming bearing which had always characterized him remained with him to the end."

In a memorial address delivered by Professor Herman Ebert before the Physical Society of Erlangen on March 7, 1894, the following sentiments are expressed: "In him there passed away not only a man of great learning, but also a noble man, who had the singular good fortune to find many admirers, but none to hate or envy him; those who came into personal contact with him were struck by his modesty and charmed by his amiability. He was a true friend to his friends, a respected teacher to his students, who had begun to gather round him in somewhat large numbers, some of them coming from great distances; and to his family he was a loving husband and father."

It can be said in retrospect that the fundamental invention in radiotelegraphy was made by Hertz, and yet it is true that the discoverer of electric waves had no anticipations as to their utilitarian possibilities. There was no rush to the patent office; indeed, it was not until two years after Hertz's death that the first application for a radio patent was filed, by Marconi. The chief interest at the time was purely scientific, the results being hailed as the settlement of a great scientific controversy, the confirmation of Maxwell's theory, the annexation to electricity of the entire domain of light and radiant heat. In the current literature we find little of prophecy with respect to utility. Sir William Crookes has been credited with being one of the first to foresee distinctly the applicability of "Hertzian" waves to practical telegraphy. In an article in the *Fortnightly Review* for February, 1892, he made a remarkably accurate forecast of what was to come: "simpler and more certain means of generating electrical rays of any desired wave-length"; "more delicate receivers which will respond to wave-lengths between certain defined limits and be silent to all others"; "means of darting the sheaf of rays in any desired direction. . . ." And for secrecy he foresaw that "the rays could be concentrated with more or less exactness on the receiver," if the sender and receiver were stationary; or, if moving about, "the correspondents must attune their instruments to a definite wave-length. . . ." "This is no mere dream of a visionary philosopher," he wrote. "All the requisites needed to bring it within the grasp of daily life are well within the possibilities of discovery, and are so reasonable and so clearly in the path of researches which are now being actively prosecuted in every capital of Europe that we may any day expect to hear that they have emerged from the realms of speculation into those of sober fact."

As we well know, all that he predicted, and more, has become reality, although progress was not to be as rapid as then seemed probable. One of those who at the time had the imagination to see, if only hazily perhaps, the great possibilities of Hertz's discovery was the youthful Marconi, who had also the initiative and the determination to put his ideas into execution, to make the new-found waves useful to mankind. Within a few years, around the turn of the new century, the world was to be thrilled by the detection of a wireless signal transmitted across the wide Atlantic. But there were insurmountable limitations to the means at hand, and it remained for still another wave in the onward roll of science, the advent of the magical era of electronics, to yield the tools necessary for the really great advance that was ahead. With the invention and development of the amplifying and oscillating vacuum tube progress was greatly accelerated, and in a comparatively short time there had been achieved, by a veritable army of experimenters, the marvels of world-wide intercommunication which are so familiar to us today.

Instruments for the New Telephone Sets *

By W. C. JONES

Transmitters and receivers for use at subscribers' telephone stations have been designed which not only materially improve transmission but also simplify manufacture and facilitate maintenance. This paper discusses these improvements and describes some of the new design technique employed in their development.

AS a result of continuous development work on transmitters and receivers for use at subscribers' telephone stations, new instruments have been designed which not only materially improve transmission but also embody features which simplify manufacture and facilitate maintenance. These instruments are now being produced for use in handsets, desk stands and wall sets in the Bell System.[1]

In many respects these instruments represent outstanding advances in transmission instrument design and performance. It is the purpose of this paper to discuss these improvements and to describe some of the new design technique employed in their development. The data presented will be confined almost entirely to physical measurements which serve to define the performance characteristics of the instruments. The interpretation of these data in terms of their relationship to the characteristics of associated apparatus and their overall reaction on transmission in the telephone plant is covered by a companion paper dealing with the transmission features of the new sets.[2]

HANDSET APPLICATIONS

The new transmitter unit with an adapter was first introduced in 1934 as a replacement for the earlier type of handset transmitter.[3] There are now about five million of these transmitters in use in the plant of the Bell System. While experience has shown that they effect an outstanding improvement in performance they do not take full advantage of the possibilities of the unit type of construction from the standpoint of simplification, owing to the fact that a number of additional parts are required to mount the unit on the existing type of handset handle. The advantages of the unit type of instrument have been realized in a new design of handset introduced during 1937, about a million of which have been produced. This handset is shown

---

* Presented at A.I.E.E. Summer Convention, Washington, D. C., June 21, 1938.

with the new combined set in the photograph, Fig. 1, and in cross-section on Fig. 2.

In designing this handset every effort has been made to obtain the maximum degree of simplicity consistent with the electrical requirements involved and at the same time to secure an attractive design which harmonizes with the other station apparatus on the subscriber's premises. Only three phenol plastic parts are employed; namely, the handle and the transmitter and receiver caps. In designing these parts particular attention has been paid to providing adequate cross-sections at the points of maximum stress and to distributing the weight so as to reduce to a minimum the breaking moments which are developed when the handset is dropped. The transmitter and receiver caps serve the dual purpose of holding the units in place and providing mechanical protection. In addition they thoroughly insulate the user from all the metal parts which are included in the electrical circuit. Both caps have smooth surfaces which can be readily cleaned. As will be pointed out later, the grid of the receiver cap also has a transmission function and plays an important part in determining the response in the upper frequency range. Spring contacts are provided to facilitate the assembly of the units in the handle. This operation is further facilitated by the fact that specific alignment of the units

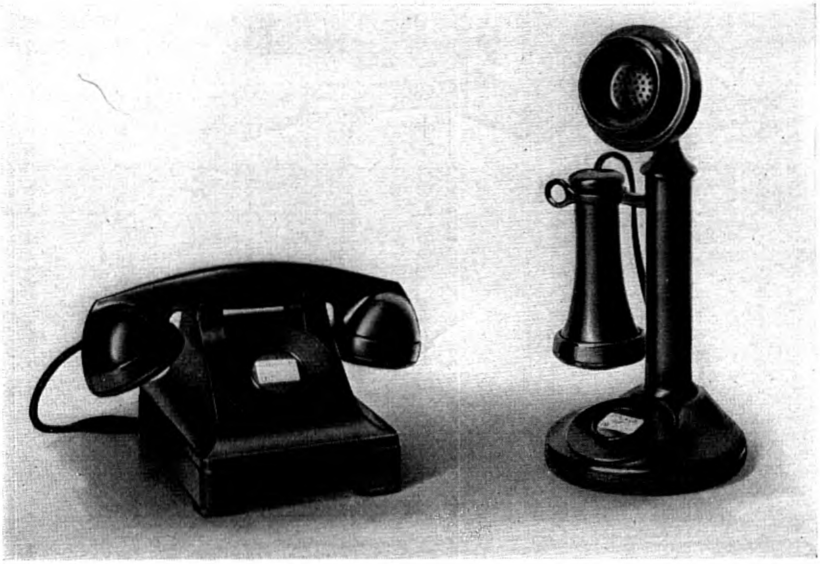


Fig. 1—Handset and desk stand equipped with the new instruments.

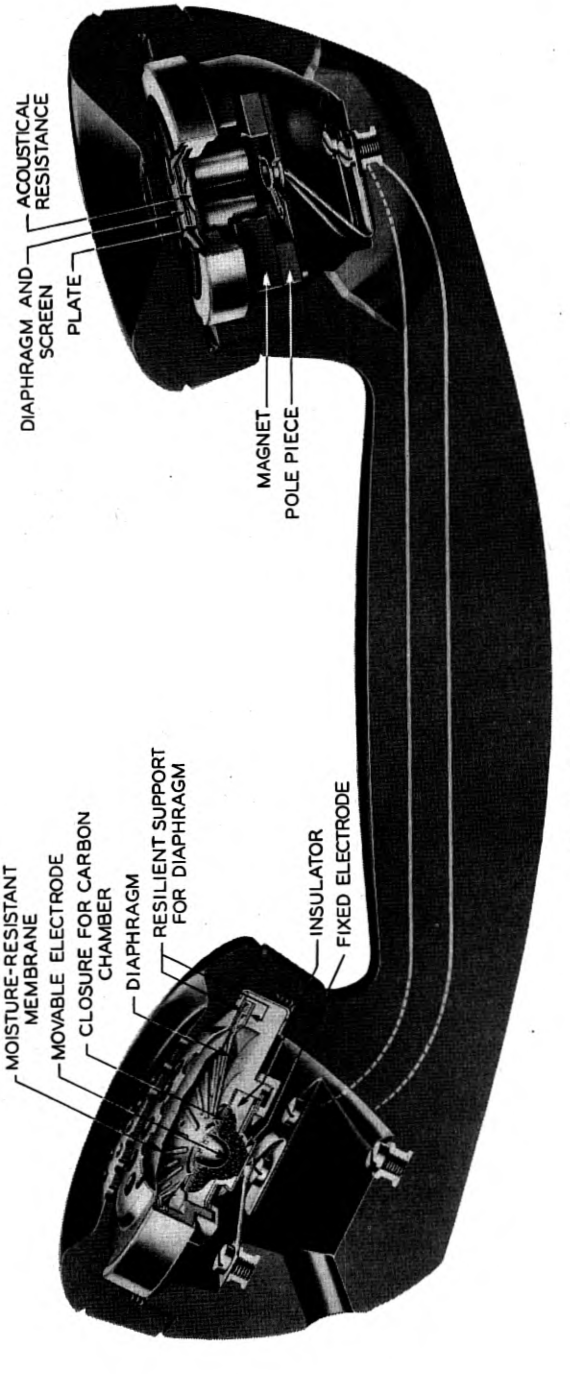


Fig. 2—Cross-section of the handset.

and the caps relative to the handle is unnecessary. The spacing between the transmitter and receiver is such that the handset can be used with the existing type of desk mounting as well as with the new combined set.

Desk Stand and Wall Set Applications

The photograph, Fig. 1, also shows the new transmitter and receiver unit adapted to desk stand and wall set use. Cross-sections of these instruments are shown on Fig. 3. The faceplate, mouthpiece and protective grid of the transmitter are combined in one phenol plastic part which is so designed as to reduce cavity resonance to a minimum and provide response characteristics essentially the same as those of the handset transmitter. On the other hand, the mouthpiece is sufficiently

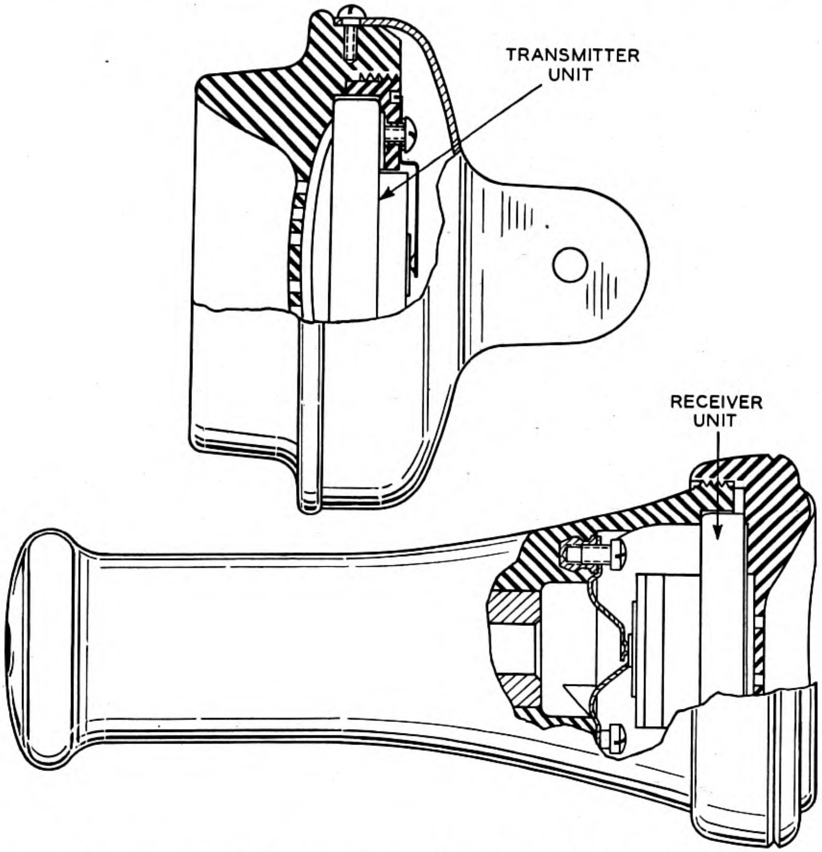


Fig. 3—Cross-sections of the transmitter and receiver for desk stands and wall sets.

prominent to encourage the user to talk directly into it and in this way reduce the losses which often result when flush type faceplates are employed with desk stand and wall set instruments. A phenol plastic part, equipped with contact springs, holds the unit tightly in the faceplate and provides electrical connections.

As in the handset the unit of the receiver is held in place by the cap. Springs are provided in the shell for bringing out the electrical connections. A metal insert adds sufficient weight to meet the switchhook requirements of the existing sets. The phenol plastic parts of both the receiver and transmitter are so designed as to insulate thoroughly the units and minimize breakage.

TRANSMITTER UNIT

The new transmitter unit is of the "direct action" type, that is, one in which the movable electrode serves the dual purpose of contact and pressure surface. As is shown by Fig. 2, this electrode is mounted at the center of a diaphragm of thin aluminum alloy formed into a shallow cone and ribbed to add rigidity. "Books" of thin impregnated paper mounted in a recess in a die-cast frame provide a resilient support for the edge of the diaphragm. The fixed electrode is held in place in the frame by a threaded ring and is insulated from the frame by a phenol fibre washer and a ceramic insulator which also forms one of the surfaces of the carbon chamber. The active surfaces of both electrodes are gold plated. A silk annulus clamped at its outer edge between the ceramic insulator and the frame and its inner edge between the movable electrode and the diaphragm forms a resilient closure for the carbon chamber. Electrical connection between the movable electrode and the frame is provided by means of metal strips of low stiffness. Provision is made for machine filling the carbon chamber through a hole in the fixed electrode and closing this hole by means of a cap which crimps over a projecting shoulder. The exposed surfaces of the cap and the threaded ring are silver plated and form the contact surfaces for the electrical connections. A moisture-resistant membrane protects the internal parts of the unit from the effects of condensed moisture from the breath. This membrane is clamped at its outer edge between a protective grid of perforated metal and the frame. A thin metal ferrule fastens the grid to the frame. The exposed parts of the unit are anodically finished to resist corrosion.

In addition to being simpler than the earlier transmitter and hence less difficult to produce, the new transmitter unit has characteristics such that:

1. Its performance is less affected by angular position.
2. There is less aging under the conditions encountered in service.
3. The electrical output is higher and the response more uniform.
4. The modulation products resulting from non-linearity are materially reduced.

Effect of Angular Position.—In order to insure good contact between the carbon granules and the diaphragm in the positions in which the handset is most likely to be held in service, the carbon chamber of the earlier transmitter was placed in front instead of the conventional location in back of the diaphragm.[3] The positional characteristics of this transmitter were further improved by the use of a "barrier" type of variable resistance element in which the electrodes are stationary and form the walls of the carbon chamber, and in which the surface of the diaphragm in contact with the granules is insulated and serves only as means for changing the contact forces between the granules in response to the variations in sound pressure at the diaphragm surface. While this transmitter represented a distinct advance in handset performance from a transmission standpoint and was quite effective in reducing undesirable positional effects, particularly in the "horizontal face-up" position, it was somewhat complicated mechanically and involved the problem of providing a closure between the diaphragm and the adjacent electrode which would be sufficiently resilient to meet the transmission requirements and at the same time prevent carbon leakage. In addition, there was some degradation in quality when it was held in the "horizontal face-down" position where the carbon granules tended to fall away from the diaphragm. While this condition occurred only infrequently in service, it was one which it was considered desirable to eliminate if this could be accomplished without making the structure mechanically complex or difficult to manufacture or maintain. A tendency also was observed in the field for the resistance to increase sufficiently under certain conditions to react adversely on the operation of the associated signaling apparatus. Owing to the inherently small areas of the sound passages leading to the diaphragm the moisture condensed from the breath could not be excluded by a membrane without complicating the structure and adding sufficient mechanical impedance to impair transmission.

Following the introductory work on the barrier transmitter, an intensive study of the direct action type of carbon element was made to determine whether the limitations of the earlier structures of this type, which arose from the non-fluid character of the carbon, could be overcome. This study resulted in the transmitter unit shown on Fig. 2. This unit eliminates the undesirable features of the inverted type without sacrificing its desirable characteristics.

The electrode surfaces of the new transmitter unit are so proportioned and so spaced relative to each other that the important current paths shift their locations in the carbon mass with changes in angular position in a manner such that the mean effective pressures in the paths and the lengths of the paths result in substantially constant resistance in all positions. Furthermore, the components of the axial motion of the diaphragm effective in changing the contact forces in the paths are also such as to produce essentially constant modulation. Not only is the total resistance of the paths between the electrodes substantially constant, but this resistance also is uniformly distributed between the individual contacts with the result that at no time does the contact potential rise to a value sufficiently high to produce objectionable carbon noise or "burning." These features result in resistance, volume efficiency and carbon noise characteristics which, as is shown by Fig. 4, are essentially independent of angular position.

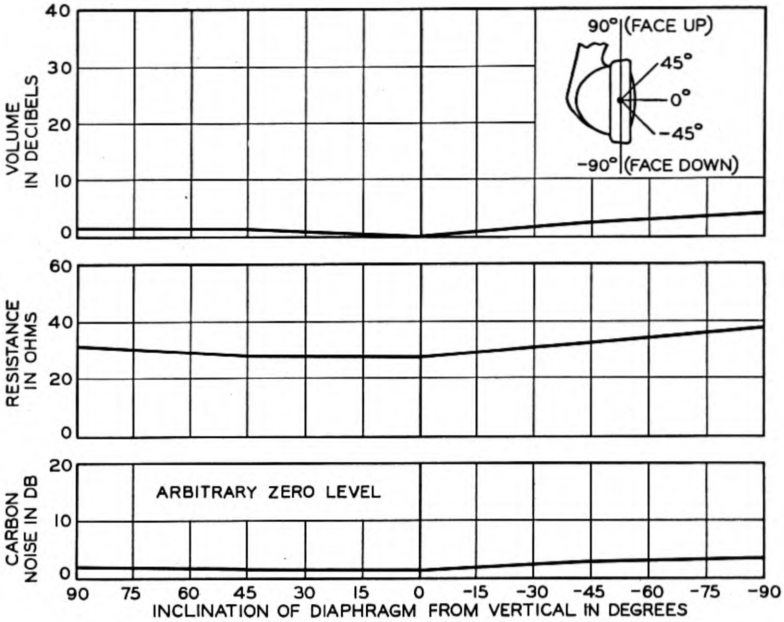


Fig. 4—Positional characteristics of the transmitter.

Another and perhaps a more exacting criterion of the adequacy of a transmitter from the standpoint of its ability to function satisfactorily in all positions is the extent to which its transmission characteristics at normal speech intensities are adversely affected when immediately preceded by loud speech. If a poorly designed transmitter is held in a position such that the carbon granules tend to fall away from the movable electrode when this test is applied, the non-fluid action of the carbon will prevent the reestablishment of contact with the electrode surface with the result that volume losses of as much as 20 db and a serious degradation in quality take place. Furthermore, these losses persist until the transmitter is jarred or moved about sufficiently to change the configuration of the granules. On the other hand, if the effect of the frictional forces within the granular mass has been taken fully into account in the design of the carbon element, these forces will not react in a manner such as to prevent good contact with the electrode following the large amplitude produced by loud speech and uniform volume and good quality will obtain at all times. The new transmitter meets this test with a substantial margin.

Carbon leakage is prevented in the new instrument without impairment of transmission by the resilient silk closure for the carbon container previously mentioned.

Aging.—Transmitter design has advanced to a stage where heating at the points of contact in the carbon element need no longer be an important source of aging. Therefore, such aging of the granular material as occurs in a well-designed instrument is limited almost entirely to that resulting from changes in the properties of the granules caused by abrasion of their surfaces when the transmitter is subjected to mechanical shocks such as occur when the handset is placed on the mounting. As in the case of the earlier transmitter, the new transmitter is machine filled[3] with the result that the motion of the granules and the resultant surface abrasion is reduced to a minimum.

The changes in resistance due to the residual aging have little adverse effect insofar as volume is concerned. In fact, the constants of most of the circuits in which the transmitter is used are such that an increase in resistance adds to rather than decreases the electrical output because of the greater amount of power supplied to the transmitter from the central office battery.

On the other hand, an increase in resistance, though small, may prove to be important in certain circuits where a critical relationship between transmitter resistance and the performance of associated apparatus exists. Under these conditions variations in transmitter resistance may result in failure of the associated apparatus to perform satisfactorily if certain limiting values of resistance are exceeded. In determining the limits to be placed on these values account must be taken of all the variables in the circuit in which the transmitter is connected. Obviously combinations of variables of this nature cannot be dealt with on the basis of averages alone but must include some measure of their range, such, for example, as the standard deviation.[4] The available data indicate that average transmitter resistances and standard deviations which lie within the area bounded by the dotted curve, Fig. 5, will have no adverse effect on circuit operation in the Bell System plant. This curve is based on certain marginal circuits of which there are a number in everyday use. An important transmitter resistance in determining the performance of associated apparatus in these circuits is the resistance during the period when the call is being established. This resistance is referred

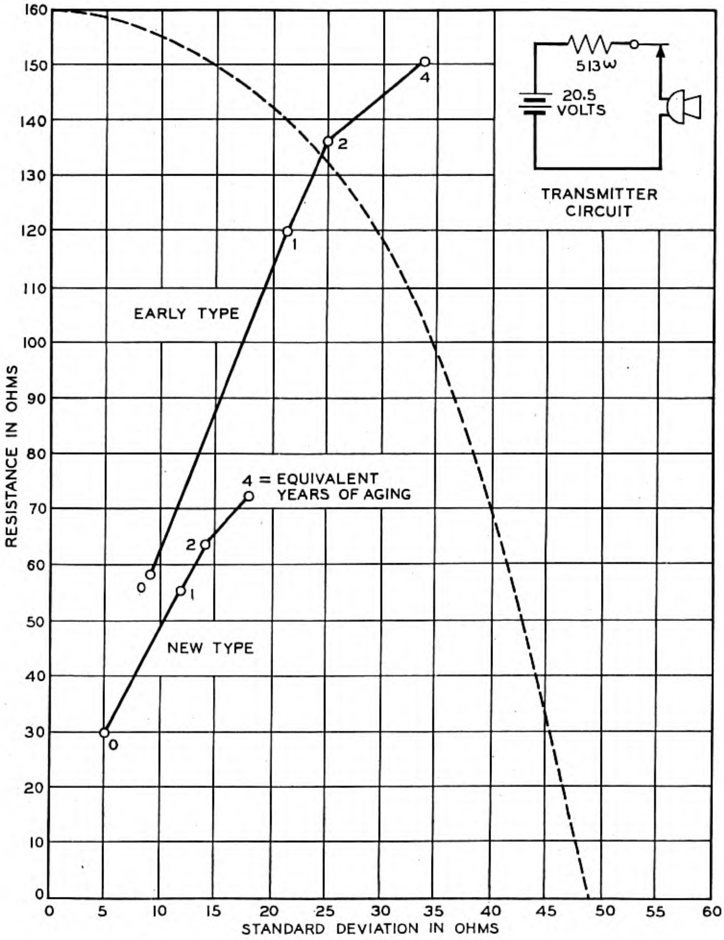


Fig. 5—Limiting values of transmitter resistance.

to as the signaling resistance. As is shown by the solid curves the signaling resistance of the earlier type of transmitter after artificial aging by an amount considered to be the equivalent of four years in service falls outside the acceptable area. On the other hand, the resistance of the new type when aged and measured under identical conditions falls well within the limiting curve and hence not only requires less frequent replacement but also permits greater freedom in circuit design and plant layout.

Moisture condensed from the breath is an important factor in determining the life of a transmitter. A protective membrane is provided in the new transmitter unit which not only is highly moisture resistant but also results in no appreciable transmission impairment. The characteristics of the material employed in this membrane are such that it is not affected by the aging conditions encountered in service such, for example, as the alkaline reaction of water after it has been in contact with phenol plastic parts or tobacco ashes. The exposed metal parts are finished to resist the corrosive action of these agents.

Response.—Reducing the transmitter to an equivalent electrical circuit provides a useful means for analyzing its performance and determining the extent to which the individual parts contribute to its response. Such a circuit for the new unit is shown on Fig. 6.

While the diaphragm can be represented as a lumped mass for frequencies in the region below 3500 cycles per second, it is necessary to consider it as being composed of three masses coupled by stiffnesses in order to represent adequately its performance at higher frequencies. These masses consist of the central portion m_5, the ribbed intermediate portion m_2 and the outer portion m_4. The central portion includes the mass of the movable electrode and is coupled to the ribbed portion by the stiffness s_6 which in turn is coupled to the outer portion by the stiffness s_2. The paper books which support the edge of the diaphragm have a stiffness s_4 and a resistance r_4. Their mass is included in the mass of the outer portion of the diaphragm m_4. The internal resistances of the portions which form the coupling stiffnesses s_2 and s_6 are represented by r_2 and r_6 respectively. A hole is provided in the diaphragm to permit rapid equalization of low frequency pressures of high intensity and prevent damage to the diaphragm and other parts. The mass and resistance of this hole, m_3r_3, are so chosen that their effect on response is confined to frequencies below 300 cycles per second where the station circuit itself is relatively inefficient. The controlling stiffness, s_3, is that of the cavity between the diaphragm and the die-cast frame. As is to be expected the impedance of the carbon granules is a function of amplitude and frequency. However, for the purpose

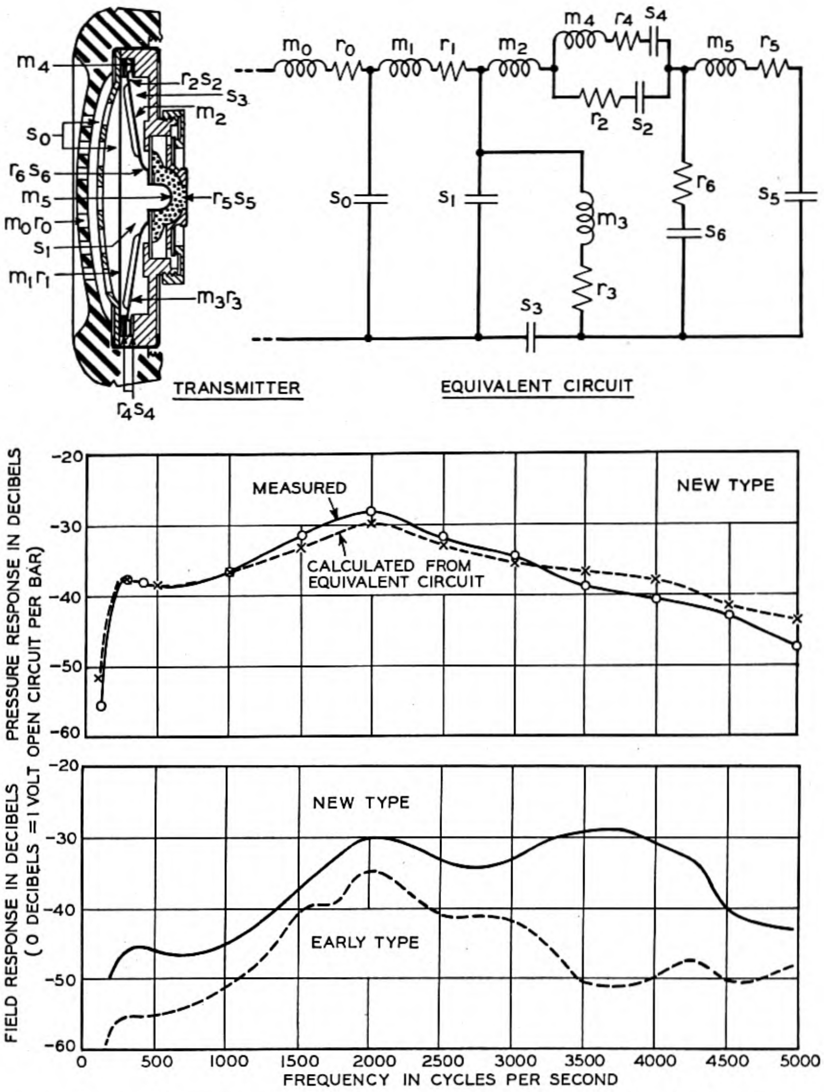


Fig. 6—Pressure and field response characteristics of the transmitter.

of this type of analysis, their impedance characteristics can be represented to a first approximation by a constant stiffness and resistance, r_5s_5. The mass of the carbon is lumped with that of the central portion of the diaphragm. The grid of the transmitter unit proper is provided for mechanical protection only and has holes large enough to have no reaction on response. When assembled in a handset or desk stand a second grid of insulating material is added. The holes of this grid have a mass and resistance, m_0r_0, which must be taken into account in arriving at an overall picture of the factors affecting response. These holes are coupled to the moisture-resistant membrane, m_1r_1, by means of the stiffness s_0 of the enclosed cavity. The cavity stiffness, s_1, couples the membrane to the diaphragm.

There are two types of response frequency measurements in general use; namely, pressure response measurements in which a constant sound pressure is maintained at the face of the transmitter throughout the frequency range covered by the test, and field response measurements in which a sound field of constant intensity is established at each frequency before inserting the test transmitter. Pressure response is used principally for purposes of analysis whereas field response usually affords a better measure of the performance of the transmitter under the conditions of actual use.

The pressure response of the new transmitter measured with a constant sound pressure at the grid, and the response computed from the equivalent circuit are shown on Fig. 6. The transmitter used in this test was artificially aged by an amount equivalent to two years of service in order to simulate more nearly plant conditions. While the computed curve departs slightly from the measured curve at certain frequencies, due to the inadequacy of some of the basic assumptions, such as those which were made relative to the impedance of the granular carbon, the agreement in general is so good as to provide a powerful tool for predetermining the response characteristics of transmitters under development and a useful method for evaluating the reaction of one element of the transmitter on another. Reducing the transmitter to an equivalent electrical circuit also has proved invaluable in determining the causes of variations in transmitter performance observed during manufacture.

As previously mentioned, the response characteristics of a transmitter under conditions of actual use may differ from those obtained with a constant pressure at its face. In general this is due to the fact that the diffraction effect of the transmitter as an obstacle in the sound field has not been taken into account. While not as readily predetermined as the pressure response, the field response can be easily measured by inserting the transmitter in a sound field of constant intensity. An artificial voice[5] is used in this measurement as a sound source and the electrical input is adjusted to maintain a constant sound pressure at the guard ring at each frequency before inserting the instrument under test. Response curves for the new and early types of transmitter obtained in this way also are shown in Fig. 6.

Both show rising characteristics in the region of 2000 cycles per second. While it is feasible to design a transmitter having a substantially flat field characteristic, experience has shown that the transmission obtained with a transmitter having a rising response in this frequency region more nearly approaches direct speech,[2] when used with a representative line and the new receiver, than that obtained with one having a more uniform field response.

Non-Linear Distortion.—A substantial reduction in non-linearity has been effected in the new transmitter unit.

It is a well-known fact that the slope of the input-output curve of most transmitters is not unity even for sound intensities other than those at which overloading occurs. However, this departure from non-linearity does not entirely account for the modulation products developed when two frequencies are impressed on the transmitter. Measurements of the sum and difference components in the output of the new transmitter and the earlier transmitter for frequencies of 1500 and 1700 cycles per second are shown on Fig. 7. It will be noted that whereas the difference component produced by the earlier instrument is equal to the fundamental within the speech range, the sum and difference components are both 10 db or more below the fundamental at all the intensities measured in the case of the new unit.

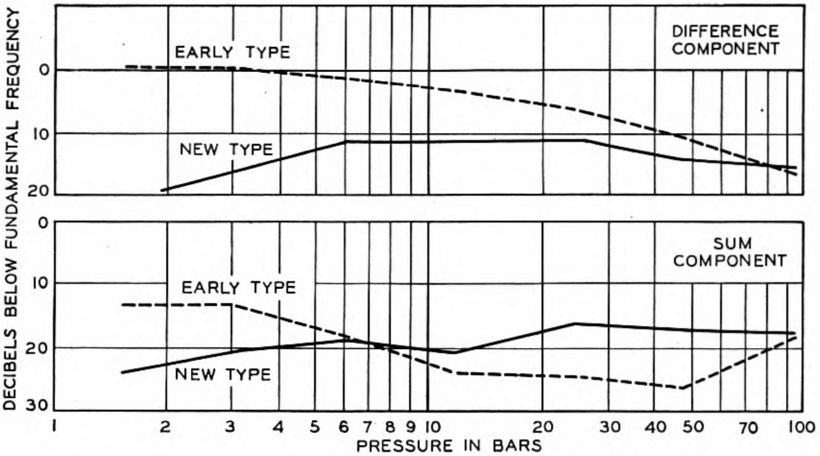


Fig. 7—Sum and difference components as a function of the intensity of the fundamental frequencies.

A tendency for the difference component to be considerably more pronounced than the sum component in the case of the earlier transmitter is characteristic of all of the commercial instruments measured during the development of the new unit. As previously mentioned, this cannot be fully explained by a non-linear relationship between input and output. The available data indicate that it is due to the manner in which the resistance changes cyclically when pressure waves of two frequencies are impressed simultaneously on the transmitter. Similar effects also have been observed in the microphonic action of carbon contacts themselves. Hence it is not unlikely that this is a fundamental characteristic of the carbon itself. If this proves to be true the extent to which an improvement can be effected in the performance of the transmitter will depend upon the ultimate control which can be exerted over the basic properties of the granular material.

RECEIVER UNIT

The new receiver unit is of the bipolar permanent magnet type. The magnetic circuit consists of pole-pieces of 45 per cent permalloy, two straight bar magnets of remalloy and a permendur diaphragm.[6] The magnets are welded to the pole-pieces to form a unit which is mounted on projecting lugs on the die-cast frame. The coils are wound with enamel insulated wire interleaved with cellulose acetate. The pole tips project through a phenol fibre plate which is fastened at the edge to the frame to form a cavity in back of the diaphragm. This cavity is connected to the recess in the handset handle or receiver shell by a hole in the plate. A disc of specially prepared silk covers this hole and provides the required amount of acoustical resistance. The silk fabric is so woven that it does not change in resistance with wetting and drying. The front of the unit is protected by a perforated metal grid which is assembled to the frame by means of a thin ferrule. A disc of impregnated silk is mounted between the grid and the frame to form a screen which prevents the transfer of foreign material from the front to the back of the diaphragm when the receiver is dropped. The grid and ferrule are anodically finished to resist corrosion. The spring contact surfaces are silver plated.

Prior to the introduction of this unit the receivers in general use for telephone purposes in this country and abroad employed simple resonant diaphragms and as a result had response characteristics which were characterized by prominent resonance peaks. As a rule the peak due to the first overtone, as well as that due to the fundamental resonance, fell within the important frequency range. This resonance not only introduced frequency distortion but increased the intensity with which circuit disturbances such as clicks were reproduced. Furthermore, the diaphragms of these receivers were rigidly clamped between surfaces which differed in temperature coefficients of expansion and heat capacities from those of the diaphragm with the result that the performance of the receiver was erratic and at a given time was dependent upon the temperature changes to which it had been subjected.

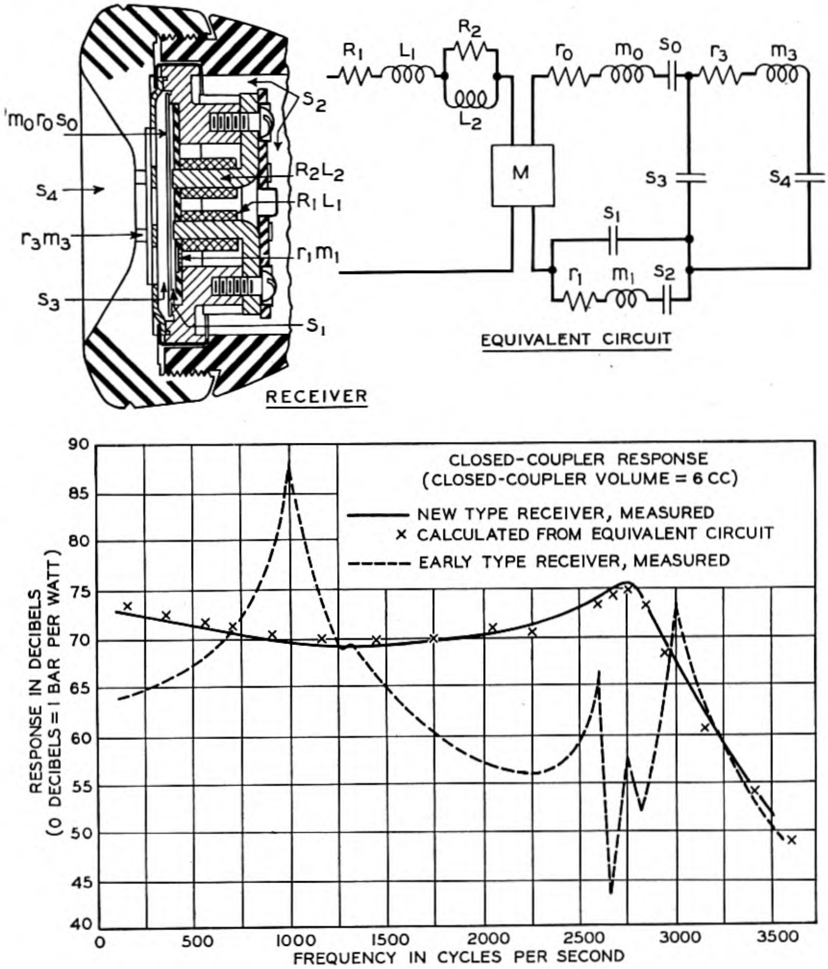


Fig. 8—Closed coupler response characteristics of the receiver.

The new receiver is so designed that:

1. All prominent resonances within the important frequency range have been eliminated and the response within this range materially improved.
2. The effect of changes in temperature has been eliminated.

3. These improvements have been accomplished without sacrificing simplicity of design or introducing features which complicate manufacture of the receiver or increase the maintenance required.

Response.—An equivalent electrical circuit for the receiver and a typical closed coupler response curve are shown on Fig. 8. Referring to this figure it will be noted that there are two meshes in the circuit which contain mass, stiffness and resistance and which control the motion of the diaphragm. One of these meshes consists of the acoustical resistance, m_1r_1, coupled to the diaphragm, $m_0s_0r_0$, by the stiffness, s_1, of the cavity between the diaphragm and the plate which surrounds the pole tips. Included in this mesh is the stiffness, s_2, of the cavity in the handset handle or receiver shell. The other mesh is composed of a cap grid, m_3r_3, and the load, s_4, coupled to the diaphragm by means of the cavity stiffness, s_3. The grid of the receiver unit proper is provided for mechanical protection only and has holes large enough to have no reaction on response. The mass of the resilient screen is small and is lumped with the diaphragm mass, m_0. The electrical portion of the circuit consisting of the winding, R_1L_1, and the equivalent eddy current circuit, R_2L_2, is coupled to the mechanical and acoustical portion by means of the force factor M[7].

The response computed from the equivalent circuit for a number of frequencies is included on Fig. 8. The agreement between this curve and the measured curve is excellent and makes it possible to predetermine the response of the receiver with a high degree of accuracy, and to evaluate the effect on the overall response of the receiver of changes in the constants of the component parts. This type of analysis also has been invaluable as an aid in establishing the causes of variations in response which have been observed during the development and production of the receiver. A measured response curve of a receiver of the earlier type has been added to Fig. 8 for convenience of reference. The improvement in uniformity and range of response is obvious. It will be noted that large gains have been effected for frequencies in the range from 1500 to 3000 cycles per second.

The response of the receiver to a square topped wave affords an excellent measure of frequency distortion. Oscillographic records of the output of typical receivers of the new and earlier types are shown on Fig. 9 for a frequency of approximately 50 cycles per second. The distorting effect of diaphragm resonance is so obvious as to require no comment beyond pointing out that for accurate reproduction of square waves uniform response for an infinite frequency range is required and that the slight rounding of the corners of the wave as reproduced by the

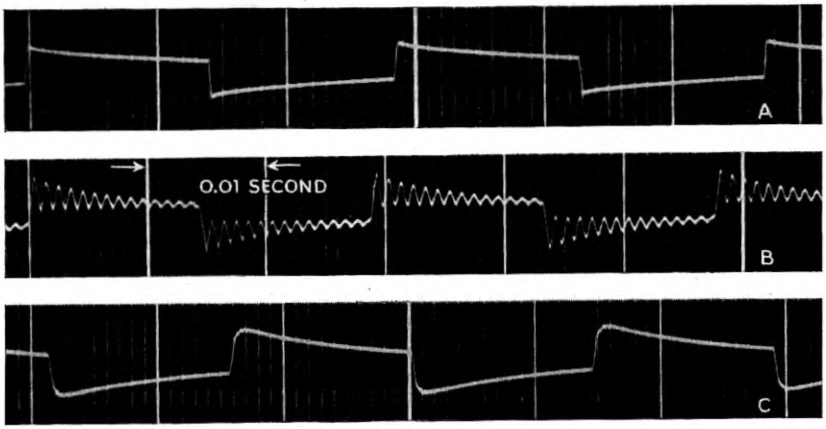


Fig. 9—Response of the receiver to square topped waves.
A—Measuring circuit—no receiver.
B—Early type receiver.
C—New type receiver.

new receiver is due primarily to the falling off of its response above 3000 cycles per second.

The substantially uniform response of the new receiver also renders clicks and other surges much less objectionable. This is due to the fact that the ear does not respond to the peak value of an oscillatory transient alone but integrates the oscillation over an interval at the beginning of the surge, hence the higher the damping the less objectionable the click.

The non-linear distortion produced by a receiver of the new type is negligible in its reaction on transmission, the harmonics in the output being 35 db or more below the fundamental.

Magnetic Circuit.—Inasmuch as the magnetic properties of the diaphragm, as well as its mechanical properties, must be considered in arriving at the preferred dimensions, it was necessary in designing the new receiver to develop criteria which could be applied in determining the optimum relationships between these factors. This study led to the use of the ratio of the force factor to the effective mass of the diaphragm for this purpose. For given magnetic materials in the pole-pieces and the diaphragm and a given air-gap length, there is a pole face area and diaphragm thickness for which this ratio is a maximum. Typical data illustrating this relationship are shown on Fig. 10. The available magnetic materials were studied using this technique and a decision was reached to use permendur in the diaphragm and 45 per cent permalloy in the pole-pieces.

There is a value of polarizing flux for which the force factor of the given magnetic circuit is a maximum. The rate at which the force factor falls off above and below this optimum value of flux is a function of the magnetic characteristics of the materials employed, the length of the air-gaps, etc. Without exception the more efficient magnetic circuits have been found to be the most critical as regards polarizing flux. Hence, if wide variations in the efficiency of the product receivers are to be avoided and serious losses due to subsequent demagnetization in service prevented, means must be provided not only for bringing the flux in each receiver to the optimum value, but also for insuring that it remain at this value during the life of the instrument. In order to accomplish this result the magnets of the new receiver are so designed as to overpolarize the magnetic circuit when they are fully magnetized. Equipment is provided for demagnetizing each receiver to its optimum flux value during the assembly process. Receivers

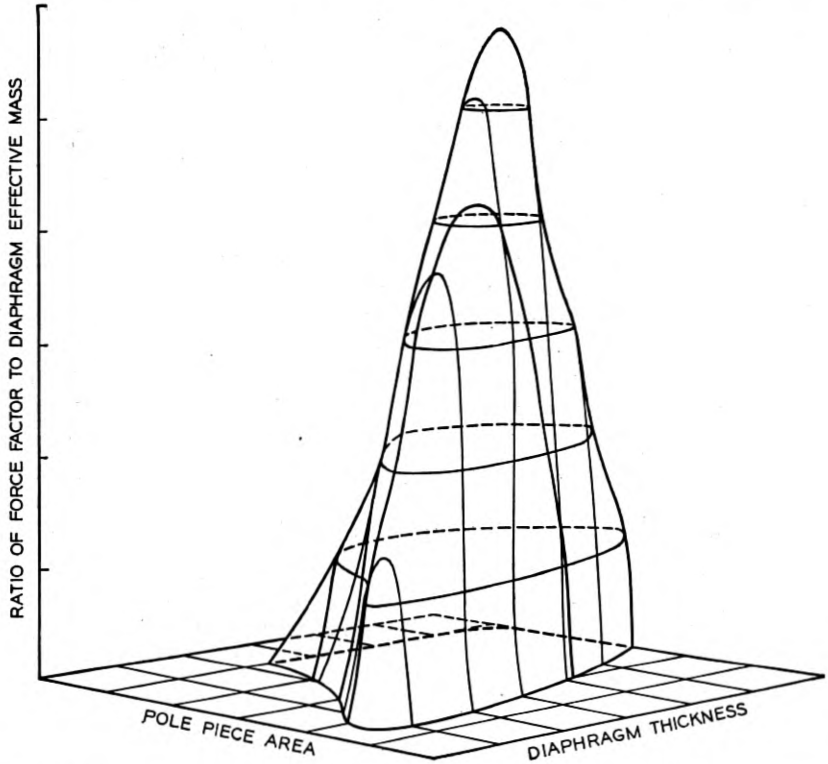


Fig. 10—Force-to-mass ratio as a function of diaphragm thickness and pole-piece area.

which are not sufficiently overpolarized before demagnetization to resist further demagnetization under service conditions are rejected.

Temperature Effects.—The diaphragm of the new receiver is held in place by the force developed by the polarizing flux and hence it is free to expand and contract independently of its seating surface. This feature renders the performance of the receiver independent of the changes in temperature to which it has been subjected. The force due to the polarizing flux is sufficiently high to prevent rattling at input intensities many times those of loud speech.

COUPLING

Although station circuits can be designed which under ideal conditions result in no sidetone, this objective is never fully realized under actual plant conditions, with the result that a part of the electrical output from the transmitter always reaches the local receiver. Whether the electrical coupling between the transmitter and receiver as evidenced by the residual sidetone is of importance from the standpoint of sustained oscillation or "howling," depends upon the degree of mechanical and acoustical coupling between the instruments. Handset and instrument design has advanced to a stage where mechanical coupling need no longer be a problem. On the other hand, as the response of the instruments is improved, the acoustical coupling may become an important item in determining the howling margin. This margin is so large under the conditions where the new handset is being used for transmission purposes that there is no tendency for oscillation or distortion to occur. However, if the handset is placed face downward on a desk or table, an air column is created which resonates in the region of 2500 cycles per second. Inasmuch as this is the region where a substantial improvement in the response of the receiver has been effected, a marked reduction in howling margin results. While there is still sufficient margin to meet all of the requirements of field use, this situation serves to emphasize the fact that such factors as acoustic coupling may limit the transmission improvements which can be effected under a given set of operating conditions.

EFFECTIVE TRANSMISSION

The extent to which the better performance of the new instruments is effective in improving the grade of transmission afforded the telephone user is a complex matter and one which is influenced by such factors as the characteristics of the circuits with which the instruments are associated at a given time, the amount of noise present at the transmitting and receiving stations, the reaction of sidetone on the loudness with which the user speaks, the distance between his lips and the face of the transmitter, the tightness with which he holds the receiver to his ear, etc. Many of these factors are beyond the control of the engineer responsible for the design of the transmitter and receiver and hence can be evaluated, insofar as their reaction on transmission is concerned, only by tests made under the conditions of actual use.

A method has been devised which makes it possible to rate the overall effect of these factors on transmission in a way representative of the results obtained by the subscribers in their normal use of the instruments.[8] Numerous tests employing this method of rating were made during the development of the new transmitter and receiver to make certain that the course followed in their development would insure the best possible performance under service conditions. Similar tests were also made of the designs selected for production. These tests show that in many respects the new instruments represent outstanding advances in transmission instrument design and performance.

REFERENCES

1. "Scientific Research Applied to the Telephone Transmitter and Receiver," E. H. Colpitts, *Bell System Technical Journal*, volume 16, July 1937, pages 251–274.
2. "Transmission Features of the New Telephone Sets," A. H. Inglis, this issue of the *Bell System Technical Journal*.
3. "Development of a Handset for Telephone Stations," W. C. Jones and A. H. Inglis, *Bell System Technical Journal*, volume 11, April 1932, pages 245–263.
4. "An Introduction to the Theory of Statistics," G. U. Yule, J. B. Lippincott Company, 1916, page 134.
5. "A Voice and Ear for Telephone Measurements," A. H. Inglis, C. H. G. Gray and R. T. Jenkins, *Bell System Technical Journal*, volume 11, April 1932, pages 293–317.
6. "Magnetic Alloys of Iron, Nickel and Cobalt in Communication Circuits," G. W. Elmen, *Electrical Engineering*, volume 54, December 1935, pages 1292–1299.
7. "Electrical Vibration Instruments," A. E. Kennelly, The Macmillan Company, 1923, page 88.
8. "Rating the Transmission Performance of Telephone Circuits," W. H. Martin, *Bell System Technical Journal*, volume 10, January 1931, pages 116–131.

Transmission Features of the New Telephone Sets *

By A. H. INGLIS

The new telephone instruments now being introduced by the Bell System result in an outstanding improvement in transmission performance in service. The evidence for this, as obtained by comprehensive laboratory and field tests, is presented here together with a discussion of the factors responsible for this superior performance and of the consideration involved in its appraisal.

NEW telephone instruments are being applied in the plant of the Bell System to the deskstand, wallset and handset, and result in markedly improved transmission performance. The new instruments are associated with the anti-sidetone feature which is also applied to the older sets already in plant. The selection of these particular designs from the wide choice made possible by new design technique, materials and manufacturing methods, has been based on developments in the methods for quantitatively rating the relative merits of different designs. In general there has been consistent effort over a period of years to base these ratings primarily on performance in service rather than on laboratory tests.

The factors influencing service performance are so many, and so complicated in their relationship, and are in so many cases difficult or even impossible for the designer to evaluate or control, that their net effect on performance cannot be predicted with certainty by laboratory methods. Of necessity such methods involve a limited selection of primary test conditions, and an even more limited selection from the possible combination of these conditions. This is particularly true in the rating of the transmission performance of a telephone set. Laboratory tests are essential in the study and analysis of design problems, and are invaluable similarly in interpolating, supplementing, and explaining service performance results. In determining the reaction on the user of the transmission features of possible designs, however, the field performance test has been found of first importance in deciding what particular characteristics to include in the new telephone instruments and circuits.

---

* Presented at A. I. E. E. Summer Convention, Washington, D. C., June 21, 1938

IMPORTANT TRANSMISSION CHARACTERISTICS OF THE NEW TELEPHONE SETS

The specific transmission design features of the new instruments are described elsewhere.[1] The purpose here, therefore, is to discuss primarily the outstanding improvements in performance resulting from the application of the new instruments and the anti-sidetone feature which has been available for some time.

These improvements are

1. Those due to the station circuit, which, as compared with the previous station circuit,
    a—largely reduce the efficiency of the sidetone path between transmitter and receiver without materially affecting the electrical efficiency of the set in transmitting or receiving. This means that sounds, either noise or speech, which are picked up by the transmitter are reproduced in the receiver of the same set at a much lower level.
    b—reduce the susceptiveness for certain types of party line sets to interference with reception by noise set up by power transmission systems.
2. Those due to the physical characteristics of the transmitter and receiver.

Fig. 1—The new handset and deskstand telephone instruments.

Several of these features have been available for some time and have, of course, been introduced into the plant as they became available. The new transmitter and the anti-sidetone circuit, for example, have been standard for some years and have already been installed in large numbers.

Figure 1 shows both the new handset and the deskstand forms of mounting, including all these features as integral parts of their design. The new desk type transmitter and receiver can, of course, be used with wall sets.

The schematic drawing of Fig. 2 indicates the general arrangement of parts in the new station transmission circuit for either type of set.

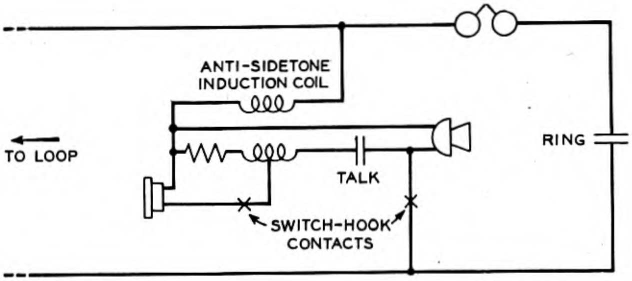


Fig. 2—Schematic transmission circuit of anti-sidetone coil.

In describing the results produced by these transmission features, and the methods employed in measuring and rating these results, it seems desirable to include some discussion of the characteristics of a telephone conversation as distinguished from a direct, face-to-face conversation, so that the various effects of the new circuits and instruments may be seen in as correct relative proportion and as generally comprehensible form as possible.

Some Elements of the Station Transmission Problem

In either a telephone or a direct conversation, successful cummunication depends on the characteristics of the talker and of the listener, and their reactions to each other and to the character of their surroundings. In a direct conversation such, for example, as across a desk, the environment is in general the same for both talker and listener, and their ears are materially aided by their eyes. In a telephone conversation, however, not only may the surroundings of talker and listener be entirely different, but a third element, the telephone system, is added to the environment of each user, which complicates his reaction, not only to his own surroundings, but also to the other party to the conversation. Furthermore, for obvious economic reasons, the natural binaural reception of direct conversation, with its advantages in discriminating between sounds from different directions, is replaced in the telephone conversation by a monaural medium.

Fundamental differences of this kind between telephone and direct conversation must be taken into account in the design of a telephone transmission system if satisfactory results are to be obtained. For example, the talker is accustomed in a direct conversation to regulate his talking volume by what he himself hears under prevailing noise conditions (which incidentally are the same for the listener), by the ease with which he hears the other party, and by the ease with which the listener appears to hear him. By experience, under ordinary conditions, the first factor mentioned, the loudness with which the talker hears himself, probably comes to be the primary control on his talking volume.

These various factors also serve to regulate talking volumes in conversation by telephone, but their magnitudes and the relations between them differ from the condition of face-to-face air path conversation and vary from one type of telephone connection to another. For example, the "sidetone" of the telephone set, being materially higher than the air-path sidetone, deceives the talker, not only by making him think he is talking louder than he really is, but also by apparently modifying the noise conditions under which he is talking in the pickup and amplification of room noise by his telephone transmitter. Since, in addition the efficiency of the telephone circuit itself may be different in the two directions of transmission, the loudness heard by one party may differ more from that heard by the other than in the case of air transmission. Then, too, noise conditions may be and frequently are quite different at the two ends of the telephone circuit. Figure 3*A* shows the probability of noise of various average intensities at subscribers' stations as determined by several surveys covering a large number of locations. On the assumption that any one of the stations represented by these data may with equal probability call any other one, Fig. 3*B* has been computed, showing the probability of noise at the two stations of a telephone connection differing by more than a certain amount. It will be noted that there is about an even chance of the noise at the two ends differing by more than 12 db. In view of these differences, a person's judgment of how well he is heard and understood can not be as direct as in the case of air transmission.

In addition, the transmission over the commercial telephone system affects the quality of the received speech more than the usual room surroundings in air-path transmission. While acoustic resonance and reverberation in a room do distort speech, in the extreme case to a point where understanding may be difficult, such a condition is distinctly unusual. Equal freedom from distortion in a telephone system is a more difficult and expensive condition to obtain than in direct conversation a few feet from a listener. Something less than perfect reproduction must suffice, for the present anyway, if costs are not to be prohibitive.

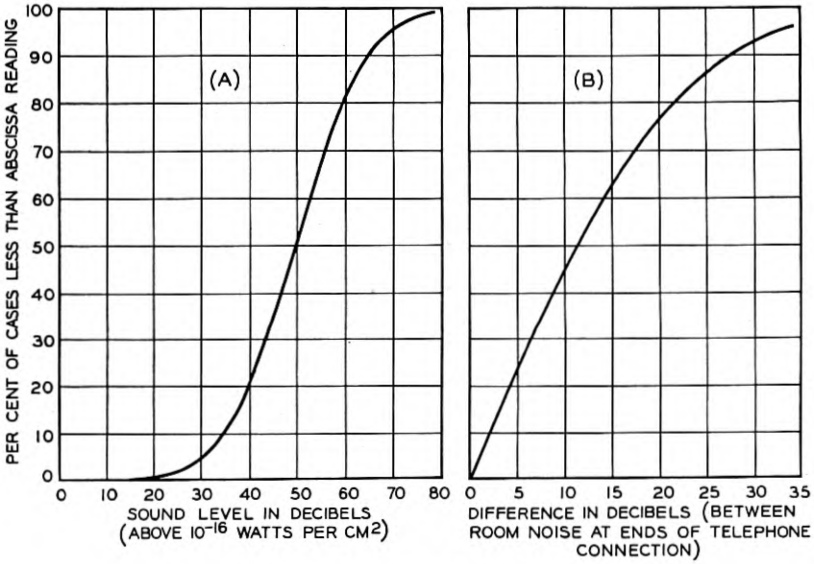


Fig. 3—Noise conditions at telephone stations.

All of these differences involve the acquiring by the user of a set of telephone habits which differ from those he has acquired in direct conversation. The problem of the transmission design of a practical telephone system requires, then, for a satisfactory solution, not only a determination of the proper speech levels to be delivered, and of the sidetone characteristics which will, under the conditions of a telephone conversation, give optimum results with the noise encountered, but also a decision as to what particular frequency range and characteristic to choose. Properly designed, a telephone transmission system should minimize, to the degree consistent with costs, its inherent differences from direct conversation, and make it easy for the ordinary user to get, without undue effort, results which are satisfactory to him in comparison with direct conversation.

In the earlier days of telephony, the problem presented appeared much simpler. It was, in effect, uni-dimensional, calling primarily for more efficient instruments and circuits; more and more power delivered to the listener's ear. While methods for the control of sidetone were not unknown, the importance of such control was not fully appreciated. Little choice was available in the quality of reproduction provided by transmitter and receiver, because of meager design knowledge.

Relatively recent and quite rapid developments in knowledge of the problems involved, in materials, methods, and measuring facilities, have now presented the necessity for a solution in essentially a three-dimensional form. These three dimensions may be described as volume, noise and quality. The solution of the problem on this basis is obviously more difficult, and has required the development of methods for quantitatively evaluating and rating their characteristics in terms of some common yardstick.

METHODS OF RATING TRANSMISSION

For reasons already suggested, such a yardstick must be based on service performance—on the results obtained by actual users in the course of day-to-day telephone service.

Extensive investigation has indicated that the best comparative measure of this transmission performance in local exchange service is to be found in the time rate of the occurrence of repetitions required by subscribers for understanding telephone conversations.[2] Or, more explicitly, when two transmission conditions have the same repetition rates, all other service factors being equal, these conditions are taken to be equal with respect to transmission performance. Where two conditions are not alike it is usually possible to evaluate the difference in the repetition rates for the same users by inserting distortionless loss in the better condition until both have equal repetition rates.

Thus, by taking as a reference a typical telephone circuit of specified make-up, the effects of various factors such as distortion, noise, attenuation, sidetone, or type of instrument, may all be expressed in the common terms of the reference circuit trunk which will give the same repetition rate.

Instead of making this adjustment in every case for the purpose of evaluating the relative performance of different test conditions for the same users, the evaluation may be made rather closely over a limited range by the following typical relation derived from repetition observations on circuits containing trunks, the losses of which were varied over a range of values.

$$\text{db} = 50 \log_{10} R_1/R_2,$$[3]

where R_1, and R_2 are the repetition rates of two conditions under comparison, and the db figure is the change in the reference trunk which has the same relative effect on the repetition rate.

Such a method is, of course, somewhat cumbersome, and requires a large amount of data to iron out random variations and individual peculiarities of little general interest. But as the fundamental rating method, supplemented by laboratory test, it has been systematically used in studying the value of the anti-sidetone circuit and in selecting instrument characteristics.

Supplementing the repetition observations, it has been found useful in service rating to obtain data on speech levels delivered to the line for each condition observed. This has been done with the volume indicator, a vacuum tube voltmeter so designed that the reading is approximately proportional to mean syllabic voltage.[4] The information thus obtained is useful not only in analyzing the results of service tests but also in determining typical values for speech levels, necessary for laboratory tests.

Laboratory tests are of two general types: objective measurements, and subjective tests. Transmission measurements cover a wide field with objectives ranging from the physical analysis and study of different designs, to the determination of overall performance characteristics of structures and systems. It is these latter tests that we are more particularly interested in here, as most descriptive of the physical properties of importance in providing telephone transmission service.

Subjective tests in the laboratory may be said to be midway between physical measurements and field performance tests. Made under controlled and somewhat artificial conditions, they indicate quantitatively the capabilities of a telephone system in transmitting articulate speech under the particular conditions of the test. They cannot, of course, indicate the relative probability of occurrence, and hence importance, of these different conditions, nor predetermine how well the subscriber will avail himself of the capabilities provided.

Consideration of some of the results of investigations in both laboratory and field will do much to explain the rather large transmission improvement realized by the introduction of the new sets in actual service, particularly if examined with the conditions of a direct conversation as a basis of comparison.

THE STATION CIRCUIT

There are two characteristics of the new station circuit of particular importance from a transmission standpoint.

Reduction of Sidetone

The first is the anti-sidetone induction coil through which the transmitter and receiver are coupled to the line. This coil comprises, in addition to three transformer windings, a balancing network. The circuit, made up of the four elements: transmitter, receiver, line, and network, coupled by the transformer, functions in such a manner that the transmitter and receiver are in conjugate relationship, i.e., voltages produced by the transmitter are balanced out and do not affect the receiver. Theoretically, such a circuit, with pure resistance elements, can be perfectly balanced at all frequencies with complete elimination of sidetone, and at the same time be as efficient as can any transformer coupling in an invariable telephone set,[5] for the transfer of power from the transmitter to the line, and from the line to the receiver.

This type of circuit is not new in principle, and many varieties are known and have been described.[6] Many of these arrangements, for one reason or another, are not suitable for application. Some, for example, call for impedances of transmitters or receivers differing widely from those available. Certain others are not economical for common battery service, where the transmitter must receive its battery supply from the line. Still others require relatively complicated and expensive cording and switchhook arrangements. The circuit which has been chosen for general common battery subscriber station application, and shown schematically in Fig. 2, is not only as simple and as easily adapted to Bell System conditions as any, but permits a coil design which is economical to manufacture as well as efficient in performance. Other types of anti-sidetone circuit have been adopted for local battery station service and for operators' telephone sets.

The theory of operation of this anti-sidetone circuit has already been discussed elsewhere.[7] It is intended here to show the general purposes of the application, some of the considerations involved in the design, and the kind of results accomplished.

While in theory complete elimination of sidetone is possible, as well as ideal efficiency of transformation, in practice neither objective can be entirely realized. The unavoidably wide variations in line impedance looking from the set, ranging from high positive to high negative phase angle, and from a few hundred to more than a thousand ohms in magnitude, together with other practical departures from ideal conditions, necessitate a choice between a high degree of sidetone balance and the standardization of a minimum number of coil designs. The variations in loop length and resistance, by their effect on transmitter battery supply, and consequently on transmitter resistance, furthermore cause variations in the absolute transmitting and sidetone efficiency of the terminal set, which must be taken into account in the station circuit design.

The actual design chosen is so arranged as to favor sidetone balance on average and shorter loop conditions where transmitter battery supply is greater, with consequent higher sidetone, and to favor transmitting and receiving efficiency on longer loops where battery supply is low. Since loop losses are greater for transmitting than for receiving because of transmitter battery supply loss, the ratio of the transformer is such as to favor the transmitting efficiency of the set somewhat in comparison to the receiving efficiency. This has the advantage of raising the transmitted speech level further above line noise. The same idea, of course, was followed in the design of the sidetone set.

The resultant anti-sidetone circuit adopted and here discussed, as compared with the sidetone circuit previously in general use, when equipped with the same transmitter and receiver and on the same loop and trunk, reduces sidetone on the average by about 10 db. Under the most unfavorable conditions of use, the reduction is unlikely to be less than about 7 db compared with the corresponding sidetone connection. Under the best conditions of balance encountered the reduction may be as much as 12 db. On the effective basis of transmission the average net improvement in transmission which results is about 6 db.

From the electrical circuit standpoint alone, the efficiency of the anti-sidetone arrangement is below that of the sidetone set in the order of about one or two db in transmitting and in receiving, which is necessitated by the limitations of practical design and circuit conditions discussed above.

Figures 4*a* and 4*b* show for transmitting and receiving, respectively, the difference in efficiency, with respect to frequency, of the anti-sidetone set from the sidetone set, each with the same instruments. Two subscriber loop and trunk conditions are shown: an average loop and trunk, and a long cable connection.

Figure 4*c* shows the variation in sidetone reduction with frequency, of the new set as compared with the corresponding standard sidetone set, for the same two circuit conditions as above. The curves are indicative of the effect of variation of circuit impedance on sidetone balance, in changing not only the magnitude, but the frequency range in which the best balance occurs.

Data of this sort alone do not, of course, indicate the relative transmission performance of the two sets. The beneficial effect on the telephone user of the large reduction in sidetone must be evaluated on the same yardstick as the losses in transmitting and receiving efficiencies which, in the practical case, accompany this reduction in sidetone. McKown and Emling have shown the effect of changes of this sort on the results obtained by the ordinary telephone user, in terms of net effective transmitting and receiving loss, as determined by service observations.[8] Their data, shown in Fig. 5, are relative to the sidetone of a reference set. The heavy solid lines are the original experimental data, the dotted extensions to these curves being extrapolated.

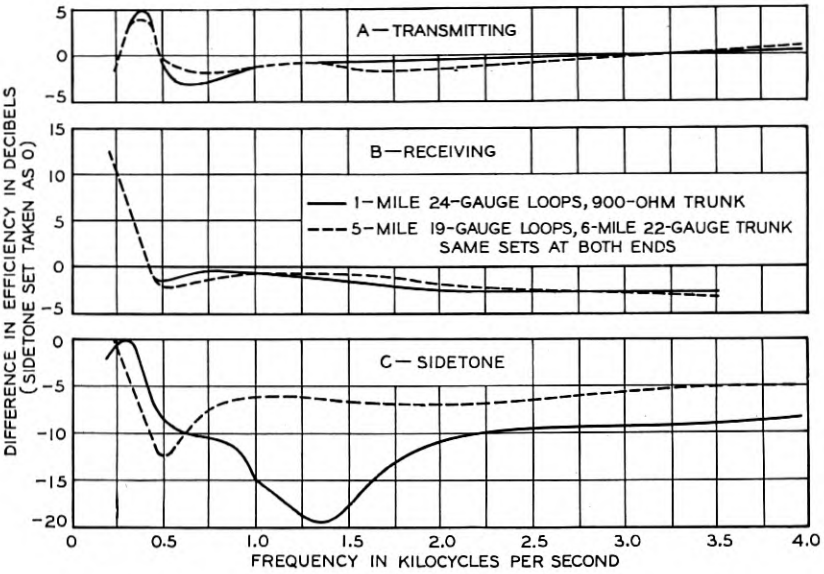


Fig. 4—Circuit efficiency of anti-sidetone circuit vs. sidetone circuit.

In addition to the original ordinates, others are shown which are of interest. They are based on the results of loudness balance tests, and while not perhaps of great precision, do approximately indicate the relationship of the sidetone of a telephone conversation to that for direct speech, and illustrate the differences in the effects of sidetone for transmitting and for receiving.

On each curve are indicated the average sidetone value of the standard sidetone and the new anti-sidetone set, each, as before, with the new transmitter and receiver. There is also shown the range of sidetone for each type of set, within which practically all service conditions will fall. This indicated range takes into account not only variations in sidetone balance due to line impedance variations, but changes (with loop resistance) in battery supply to the transmitter. It should be noted that in only a few cases is the absolute sidetone of the anti-sidetone set on the worst sidetone conditions, as high as or higher than that of the sidetone set on the best sidetone conditions, and then only by a small amount. Furthermore, in spite of the wider variations in sidetone of the anti-sidetone set, these variations are over a range such that the resultant variations in effective losses are smaller than for the sidetone set.

Considering Fig. 5*a*, it will be noted that for either sidetone or anti-sidetone sets, the sidetone is louder than for natural speech sidetone, which, as noted before makes the user think he is talking louder than he actually is. The average sidetone reduction of 10 db for the anti-sidetone set results in less of this restraint on his talking level, with a resultant net effective gain in transmitting of about 4 db compared with the sidetone set.

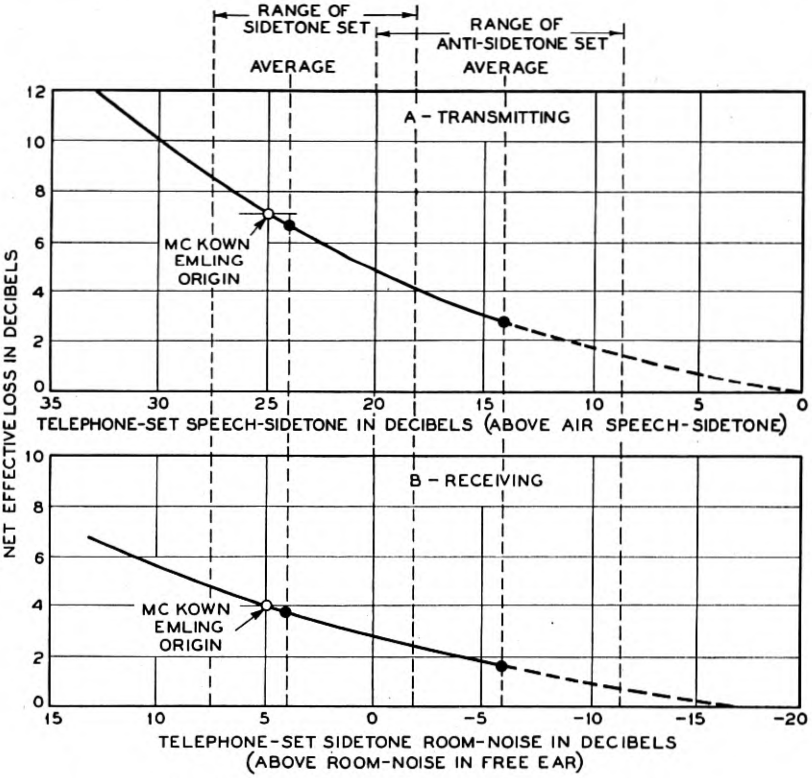


Fig. 5—Effects of sidetone on user of telephone.

In receiving, Fig. 5*b*, sidetone introduces an effective loss by the reproduction in the telephone ear of room noise picked up by the transmitter. It will be noted that for the anti-sidetone set, the reproduced noise is in general appreciably lower than the room noise itself. Inasmuch as room noise also interferes directly with received speech in the telephone ear by leakage under the receiver cap, the contribution to the total noise of the sidetone pickup of the anti-sidetone set is in most cases small. This is not true for the sidetone set, where in many cases the sidetone noise may constitute the principal interfering noise. The resultant net effective gain in receiving is about 2 db compared with the sidetone set.

This is a good illustration of the type of information which can be obtained only from a field study. For example, the relationships indicated on Fig. 5 are dependent on how far away from the mouthpiece of the transmitter and at what level the speaker talks, and on how tightly to his ear he holds the receiver. These in turn are resultants of all the conditions of the particular telephone conversation. If incoming levels are so high as to be uncomfortable, the receiver may well be held farther away from the ear. In that event, of course, the sidetone conditions of the set become relatively less controlling. The weight, size, and shape of the instrument in his hands may similarly affect the subscriber's use of it, the results he gets, and the relative importance of various factors of telephone design.

For such reasons, not only must laboratory performance tests be supplementary and subsidiary to field tests, but additional field tests must be the basis for determining the effect of any major changes in design, whether or not those changes are electrical, acoustical, or purely mechanical.

Considerations of this sort emphasize the importance of having clearly and explicitly in mind the conditions and relationships of direct conversation, as a general reference for the interpretation and explanation of the effects of telephone design on telephone conversations. The sidetone ordinates of Fig. 5, for example, not only suggest the difference in function of the anti-sidetone circuit in transmitting and receiving, but also emphasize the fact that the overall sidetone resulting from the combination of circuit, instruments, and method of use, is the important factor rather than the sidetone circuit efficiency only. Such matters are easily lost sight of, if design is not properly coordinated in its correct perspective.

The reduction of sidetone provided by the anti-sidetone sets is of further advantage in two rather different ways.

In attaching a transmitter (which is an amplifier) and a receiver, to a common handle which mechanically couples the two, a condition is set up in which the gain under certain conditions may exceed the loss in the path made up of handle, air, and electrical sidetone circuit. Sustained oscillation, or howling, will then result between transmitter and receiver. Even if this point is not reached, but is approached within 6 db or so, impairment of quality results from incipient oscillation. The greater sidetone circuit loss of the anti-sidetone circuit provides an additional margin of safety against any such condition.

The granular carbon of the transmitter, and the design of the transmitter itself must be carefully controlled, or serious noise—transmitter "burning"—will cause noise in the receiver of the set. The mechanical and electrical wear and tear of service tend to make this transmitter noise worse. In the new transmitters this "burning" has been kept at a low inherent value throughout life. The anti-sidetone circuit, however, provides a margin of safety against the small amount remaining, so that with this set there is less likelihood of transmitter noise causing impairment of reception.

Reduced Susceptiveness to Interference

It will be noted from the schematic circuit drawing Fig. 2 that two condensers are used in the new sets, one in the anti-sidetone transmission circuit, and a separate one with the ringer. In some types of party line practice the ringer of the set is connected for some parties from one side of the line, and for the others from the other side to ground.

Figure 6 shows schematically two such ringing arrangements during the conditions of conversation, *6a* as used in the new sets, and *6b* with one condenser common to transmission and ringing circuits. It will be noted that in the standard circuit adopted (Fig. *6a*), if any longitudinal noise voltages exist between the central office and station grounds, there is an equal voltage drop from each side of the line to ground through the ringing paths (assuming the two ringer condenser paths to be identical). The voltage drop across the terminals of the talking set is therefore zero and no noise results.

If the arrangement of Fig. *6b*, corresponding closely to previous designs, were used, however, this condition would not obtain. The condenser of the station in use being common to the transmission circuit as well as the ringing circuit, the noise voltage drop across this condenser is introduced in the transmission circuit. In addition there are other paths to ground from each side of the line through the transmission circuit which are not of equal impedance. The net result is a residual noise current through the receiver of the talking circuit.

In the actual case, the impedance of all ringers and condensers is not identical and there are often more parties connected to one side of the line than the other. Even under these relatively unfavorable conditions, however, the two-condenser arrangement adopted reduces the susceptiveness of the set to interfering noise by as much as 15 db. A further material reduction is realized by the high impedance of the ringer used in the new sets, so that in most cases interfering noise at grounded ringer stations will not differ materially from that at individual stations where the ringer is bridged across the line.

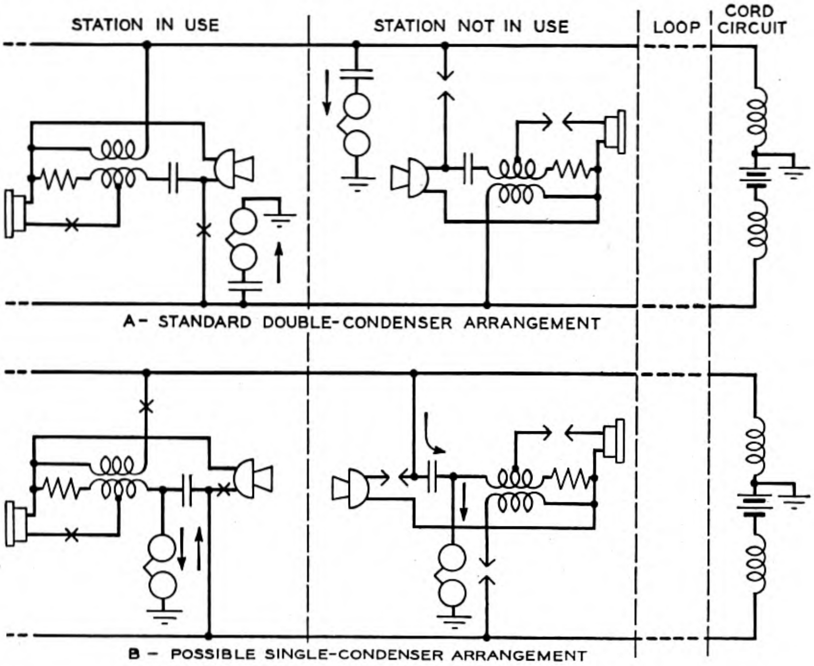


Fig. 6—Ringing arrangements for party line service.

It is interesting to note that this improvement is realized at little additional cost, since the transmission condenser, which must be of relatively high capacitance, is permanently bridged by the transmitter, so that it is protected from exposure to any large voltages, and may be of cheaper construction and smaller in size than would otherwise be the case. The ringing condenser on the other hand, while it must be constructed to withstand higher voltages, may be of relatively small capacitance, which gives more uniform and better ringing and dialing performance.

Characteristics of Transmitter and Receiver

Since the individual design characteristics of the new transmitter and receiver are discussed elsewhere,[1] attention here will be centered on the overall effects of these characteristics in the complete transmission system, as indicated both by laboratory and by field test. As stated before, the problem may be more or less arbitrarily separated into three correlated problems—volume, quality and noise.

As in the case of sidetone, these problems appear, perhaps, more nearly in a proper perspective if considered in comparison with the corresponding factors in a direct conversation. It must be remembered that telephone service does not consist in the provision of a mechanism, per se, but in the provision of facilities for conversation, to which the mechanism should be incidental, however important. Since the inherent conditions of such a conversation are quite different in many respects from those of a direct conversation with which, consciously or unconsciously, it will be compared in its overall results, the parallelism in detail should not be too exact. Departures from the conditions of direct conversation in certain respects which are relatively unavoidable, may be best compensated for by deliberate departure in certain other respects. For example, the physical absence of one party to the telephone conversation, and the monaural nature of such a conversation, may be partially compensated for by delivering to the ear of the listener a somewhat higher speech level than he is accustomed to in direct conversation. The limitation of frequency band width imposed on the telephone medium, largely for economic reasons, may be minimized in its effects if the transmission characteristics in the available band are other than a facsimile of the corresponding band in direct conversation. All such measures must be employed with knowledge of their effect on the ultimate objective, that the telephone conversation may be easy and natural.

General Requirements

It is easily seen that for any particular overall frequency characteristic of a telephone transmission system, there are practically an infinite number of ways in which it can be split up between transmitting and receiving characteristics. From this standpoint alone, then, there is no particular "best" transmitter or receiver frequency response. From other standpoints, however, certain general types of individual characteristics, both in frequency and efficiency, are to be preferred to others, particularly when considered in their practical application to an already existing telephone system. It has been pointed out [9] that in general, development has been toward a telephone system where both transmitter and receiver are relatively uniform in their frequency characteristics. Induced noise appears to be so evenly distributed with frequency that such response would not appear to magnify the interference problem.

Transmitting efficiency should be as high as required to keep the speech well above induced noise but not so high as to cause excessive crosstalk into other telephone circuits. The maximum desirable received level is determined for a given telephone system by the limitations of the human ear in accepting with comfort speech levels above a certain intensity. Finally, the practical necessity of working as satisfactorily as possible in conjunction with the telephone transmitters, receivers, and sets in the existing plant during the period of transition, places a practical limitation on the amount of change that is desirable in relative levels of either transmitting or receiving.

With regard to frequency range, previous work[10] indicated the desirability of designing circuits to transmit frequencies from 200 or 300 cycles up to about 3,000 cycles. Gains in articulation and naturalness are realized by increases in this band width, but are progressively smaller for successive equal increments in frequency. A 3,000-cycle band properly used gives good transmission both in articulation and naturalness, but frequency limitation is essentially an economic one, subject to change as conditions change. Recent work on the new multiple channel carrier systems has indicated justification in these systems for providing a somewhat wider band, from about 150 to about 3,500 cycles.[11]

Overall Frequency Response

In describing the frequency characteristic of a transmission system it has become customary to refer to it as more or less "flat," where "flat" is assumed to be synonymous with "perfect" as far as the relative transmission of various frequencies is concerned. In measurements of the elements of an electrical circuit, from which this terminology came, the word is useful since, when the measurements are properly made, at any rate, the basis of comparison implied by the word "flat" is generally understood. This is also true, although probably to a more limited extent than is generally realized, when the term is applied to electro-acoustic transmission systems, where free progressive, plane air waves of various frequencies are transferred to an electrical system, or vice versa, by means of microphones or loud speakers.

In the case of a telephone system, however, where a transmitter is placed close to the lips, and a receiver directly to one ear, and where the air waves are not free progressive, or plane, use of the word "flat" implies a basis of comparison which is not self-evident. Much effort has been given recently to establishing an appropriate reference system, sufficiently simple in concept and ease of specification, to be useful in this connection. The result of this work has been a reference telephone system which, when spoken into, would give the listener in all respects essentially the identical sensation he receives in one ear when facing the speaker directly, with an air path one meter long between the speaker's mouth and the listener's ear, in surroundings without reverberation or noise. Such a reference transmission system has tentatively been called an "orthotelephonic" system.

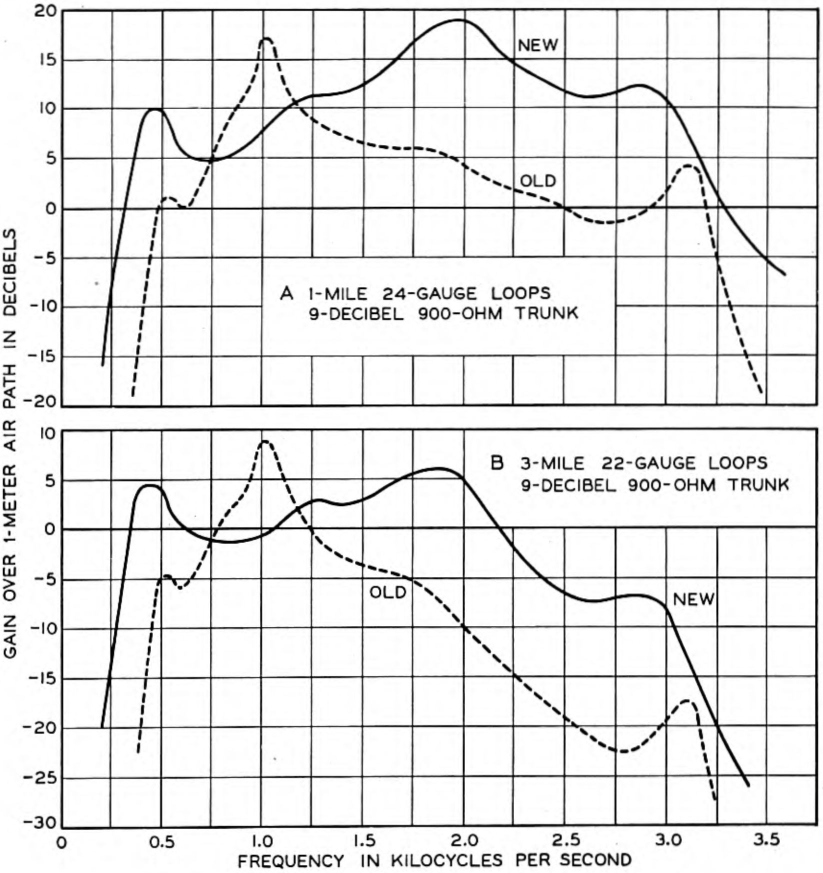


Fig. 7—Overall orthotelephonic frequency response of Typical Telephone Connections.

The point of interest here is that when measured by any suitable objective method, the frequency characteristic of this "orthotelephonic" telephone reference system is not "flat" by a considerable amount in the ordinarily accepted usage of the term. This departure from "flatness" is caused by such factors as the frequency directive characteristic of the mouth, cavity resonance of the ear, and disturbance of the sound field by the head. The individual contribution of some of these factors is not as yet definitely determined.

Furthermore, for reasons mentioned, it is not self-evident that a practical telephone system of limited frequency range should be "flat" with respect to the corresponding frequency band in this more or less basic orthotelephonic system which is not limited in frequency range. Having decided on the band width that is desirable and justifiable, it must still be determined, therefore, what particular frequency characteristics are preferable in this band.

In selecting from the many possible choices, the particular frequency response that seems best, several factors must be taken into account. This has been done by a study (under the conditions of actual service) of the relative results of several different experimental instrument designs, varying in frequency characteristics. The overall frequency characteristics of the resultant choice are indicated for two typical circuit conditions in Fig. 7. These measurements were made with the artificial mouth and ear [12] and are plotted with reference to corresponding measurements on an orthotelephonic reference telephone system. For comparison, the results of similar tests of the earlier Bell System handset [13] are shown also.

In considering these overall telephone system frequency response characteristics in the light of previous discussion, there are several points of interest:

1. The large increase in response at both higher and lower frequencies with respect to the older handset, which in itself was a notable advance in this respect over previous types. This increase amounts to 10 db or more from about 200 to 500 cycles and from about 1,700 to 3,000 cycles. This wider frequency range gives better naturalness of reproduction.
2. The type of the response. The general uniformity and absence of any marked resonance or irregularity is obvious. For either average or long loops the entire band from about 300 to over 3,000 cycles lies within a range of 15 db. It will be noted, however, that, for the average condition, the response at the higher frequencies (1,500–3,000 cycles) is distinctly above that for the frequencies below 1,500 cycles. This characteristic aids materially in the understanding of the low intensity consonant sounds. The response on the longer loops would undoubtedly be correspondingly better if the high frequencies were raised so that the overall characteristic more nearly resembled that for the average condition shown. It should be remembered,

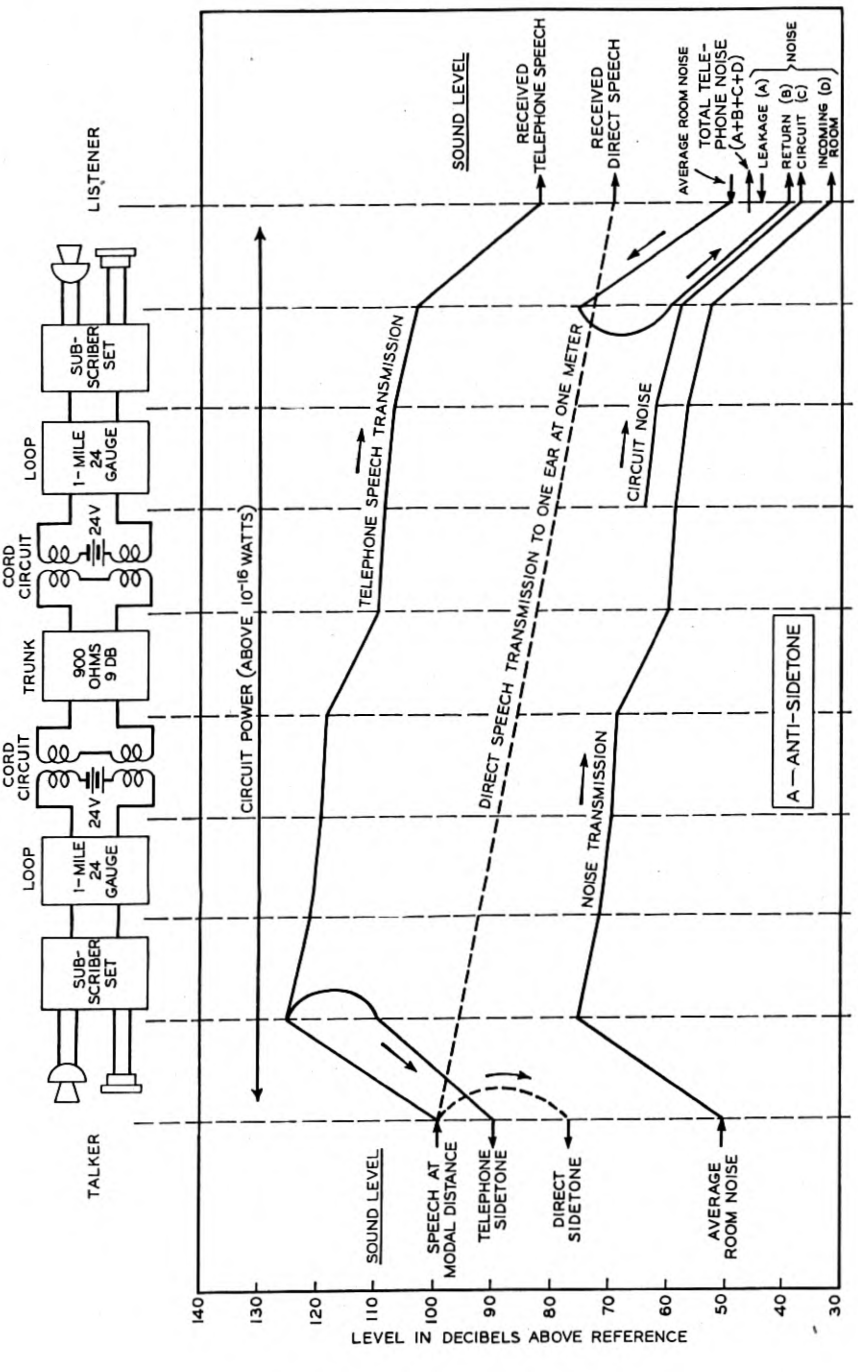


Fig. 8—Power level diagrams of typical telephone connections.

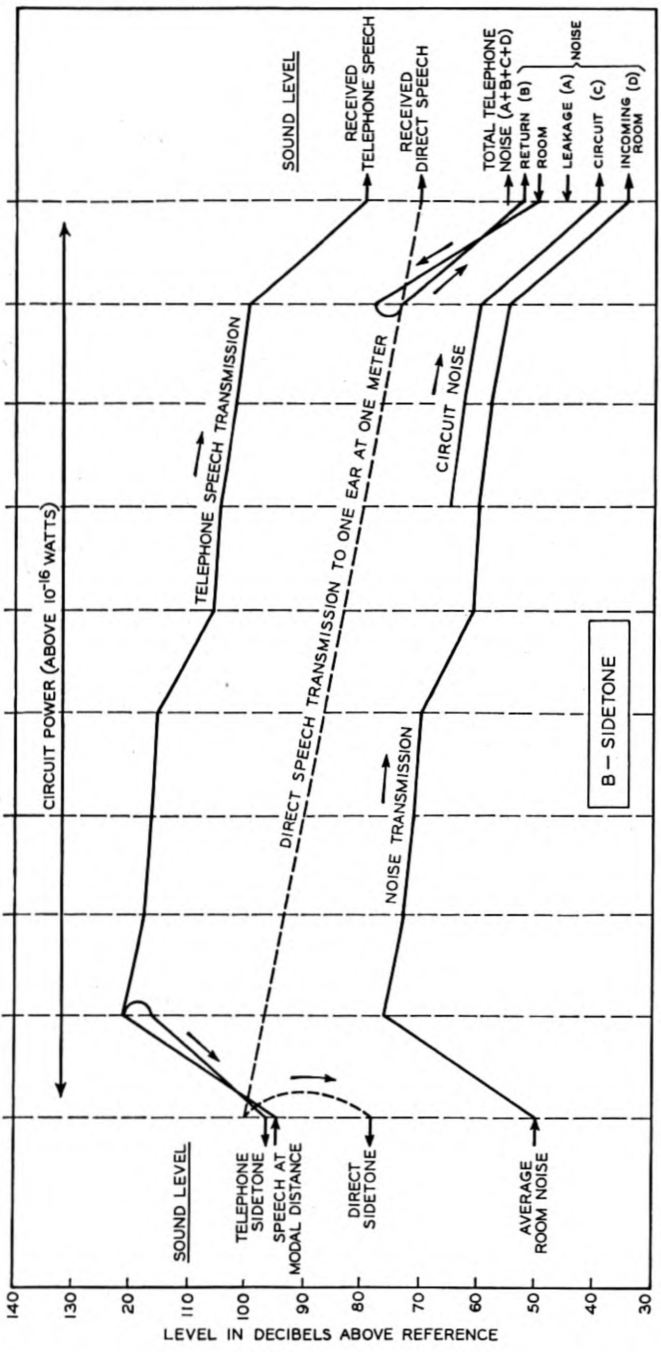


Fig. 8.—*Continued from page 376.*

however, that for a standard set, to be used on all loop conditions, a response designed solely for the long loop, with its large loss at high frequencies, would be distinctly "tinny" and disagreeable in quality on average or short loops.

3. The materially smaller losses in the transmitted band for the average telephone connection than for the orthotelephonic system. Even for the long loop conditions the losses are no greater than for this orthotelephonic system up to about 2,200 cycles. In other words, for a large majority of telephone calls the received speech level will be higher for the same talking level than in the case of direct conversation at one meter distance. The desirability of this in a monaural system of limited frequency range has already been indicated.

Combined Effects of Circuit and Instruments

The part played by the anti-sidetone circuit in permitting better utilization by the subscriber of the capabilities of the telephone system, in increasing the level of received speech, and minimizing the effect of noise is summarized by the illustrative power level diagrams in Fig. 8, using typical values of room and line noise and other conditions.

The data shown were obtained by objective measurements with the artificial mouth, sound level meter, volume indicator, and artificial ear. Frequency weightings appropriate to the levels involved were employed in the sound level measurements.

Figure 8*a* shows, for the anti-sidetone set, the losses between the talker and listener under average room noise and circuit conditions. Figure 8*b* gives corresponding information for the sidetone set. For comparison, an approximate level diagram is shown also for direct speech. The relative speech levels at the transmitter for the two telephone conditions were adjusted to give volume indicator readings on the line in accordance with the results of service observations. The upper curves in each drawing are for speech and the lower for noise picked up by the transmitters.

At the transmitting end the lower sidetone of the anti-sidetone set results in higher talking levels, with about 4 db higher speech level at the input to the circuit. In the overall circuit this gain is increased by the receiving-end effect of the anti-sidetone circuit in minimizing the effect of room noise.

The total noise in the telephone ear, as shown on the drawing, has as its principal contributing factors:

1. Leakage of noise under the receiver cap.

2. Noise picked up by the transmitter and returned via the sidetone path and receiver to the listener's ear—termed "return noise."
3. Circuit noise.
4. Room noise picked up at the far end and transmitted over the circuit.

The different relative contribution of the "return noise" for the two sets is of interest. The net result is a total noise in the telephone ear, lower than the actual room noise for the anti-sidetone set, and higher for the sidetone set.

For the circuit and room noise conditions shown, the ratio between received speech level and noise is about 25 db for the sidetone condition, and 35 db for the anti-sidetone. The corresponding ratio for air transmission to one ear under the conditions shown is in the order of 20 db.

RESULTS OF LABORATORY AND FIELD PERFORMANCE TESTS

It has been seen that the new sets are superior in volume and in minimizing the disturbances of noise. The frequency measurements just discussed have indicated marked superiority also in the quality of reproduction.

One measure of the effect of this reduced distortion is by means of the articulation test. Such tests have shown that for a typical telephone system equipped with the new telephone sets, 95 per cent of the letter sounds spoken into the transmitter are correctly understood by the listener. With air transmission, 98 or 99 per cent of letter sounds are correctly understood. The difference is almost entirely due to the broader frequency band transmitted by the air path.

With the final designs of the new sets, tests have been made by the methods developed for determining "effective" transmission.[2] The results of these tests have shown that the new sets under the conditions of actual service provide a marked advance in transmission performance. The average total transmitting and receiving gain is about 15 db on the effective basis of transmission, as compared to sets of the sidetone type used with the older type of instruments.[13]

CONCLUSION

In general it appears that the notable transmission improvement which has been achieved in the design of the new telephone sets, in their freedom from distortion, higher effective volume, lower sidetone, and general convenience in use, makes possible a closer approach to the ease of a direct conversation than has hitherto been possible commercially.

Undoubtedly, further improvements in station transmission performance will, as in the past, be forthcoming with advances in the technique of design and manufacture, and in further knowledge of the requirements of the problem.

REFERENCES

1. "Instruments for the New Telephone Sets," W. C. Jones, this issue of the *Bell System Technical Journal.*
2. "Rating the Transmission Performance of Telephone Circuits," W. H. Martin, *Bell System Technical Journal*, Vol. X, January 1931, pp. 116–131.
3. "Scientific Research Applied to the Telephone Transmitter and Receiver," E. H. Colpitts, *Bell System Technical Journal*, Vol. XVI, July 1937, pp. 251–274.
4. "Speech Power and Its Measurements," L. J. Sivian, *Bell System Technical Journal*, Vol. VIII, October 1929, pp. 646–661.
5. "Transmission Circuits for Telephonic Communication," K. S. Johnson, D. Van Nostrand Co., New York, 4th Printing, 1929, p. 105.
6. "Maximum Output Networks for Telephone Substation and Repeater Circuits," G. A. Campbell and R. M. Foster, *A. I. E. E. Transactions*, Vol. XXXIX, Pt. 1, pp. 231–280 (1920) (disc., pp. 280–290).
7. "An Explanation of the Common Battery Anti-sidetone Subscriber Set," C. O. Gibbon, *Bell System Technical Journal*, Vol. XVII, April 1938, pp. 245–257.
8. "A System of Effective Transmission Data for Rating Telephone Circuits," F. W. McKown and J. W. Emling, *Bell System Technical Journal*, Vol. XII, July 1933, pp. 331–346.
9. "Trend in Design of Telephone Transmitters and Receivers," W. H. Martin and W. F. Davidson, *Bell System Technical Journal*, Vol. IX, October 1930, pp. 622–627.
10. "Transmitted Frequency Range for Telephone Message Circuits," W. H. Martin, *Bell System Technical Journal*, Vol. IX, July 1930, pp. 483–486.
11. "Transmitted Frequency Range for Circuits in Broad Band Systems," H. A. Affel, *Bell System Technical Journal*, Vol. XVI, October 1937, pp. 487–492.
12. "A Voice and Ear for Telephone Measurements," A. H. Inglis, C. H. G. Gray and R. T. Jenkins, *Bell System Technical Journal*, Vol. XI, April 1932, pp. 293–317.
13. "Development of a Handset for Telephone Stations," W. C. Jones and A. H. Inglis, *Bell System Technical Journal*, Vol. XI, April 1932, pp. 245–263.

Spectrochemical Analysis * in Communication Research

By **BEVERLY L. CLARKE** and **A. E. RUEHLE**

The development of spectroscopy is traced through Newton, Fraunhofer and Kirchoff to Hartley, Pollock and Leonard, and De Gramont who, in the period 1880–1920, applied the spectroscope to chemical analysis. It is shown that modern quantitative spectrochemical analysis began with the comparison standard method of Meggers in 1922, and developments since that time are discussed. The organization and functioning of the spectrochemical unit of the analytical group at the Bell Telephone Laboratories is described, and a number of examples of applications to telephone problems are given.

The foundation of spectroscopy may be traced back to the year 1666 when Sir Isaac Newton discovered that sunlight is composed of the several colors, and that it could be separated into its components by refraction with a prism. Newton failed to note the discontinuous nature of the solar spectrum and it was reserved for Fraunhofer, more than a century later, to investigate the absorption lines and to point out their importance. Fraunhofer noticed that the *D* line absorption doublet in the solar spectrum was identical in position with the bright line doublet observed in a flame fed with sodium chloride. Finally in 1859 Kirchhoff formulated the modern concept of the composition of the sun based on the observed absorption lines in the solar spectrum.

As a result of Kirchhoff's theory much attention was at once turned to the examination and mapping of the emission spectra of terrestrial substances and testing for their presence in the solar atmosphere. Bunsen and Kirchhoff proved the presence of many terrestrial elements in the sun. Lockyer actually discovered one element, helium, in the sun almost thirty years before its discovery on the earth. Thus did qualitative spectral analysis enjoy a brilliant beginning.

In the years that followed Kirchhoff's work the spectroscope was properly credited with many triumphs in the field of physics, but applications to chemical analysis were extended very little beyond simple qualitative detection of the elements. In 1882, however, Hartley performed quantitative analyses by determining at what concentrations in solutions the various spectral lines of metals would disappear.

---

* Any process of chemical analysis by means of the emission spectrum will be called "spectrochemical analysis" in this paper. Absorption and fluorescence methods have also been used but these will not be considered here.

For example, in a 1 per cent solution of silver the spark spectrum contains forty-five lines, under the conditions of Hartley's experiments. In a 0.1 per cent solution of silver twenty-five lines remain; in a 0.01 per cent solution only nine lines remain; and at 0.001 per cent but one of these persists. Pollock and Leonard, following Hartley, also used solutions with an improved sparking cell and extended this method of quantitative analysis to most of the common metallic elements. Almost simultaneously De Gramont began his monumental work of determining the persistent and ultimate lines (*raies ultimes*) of all the known elements under the conditions of a condensed spark discharge.

In 1922 Meggers, Kiess and Stimson [1] published the first paper on what might be called the modern method of spectrochemical analysis. Their departure from previous methods was that they did not rely solely upon the presence or absence of certain lines as a measure of the amount of an element present in the sample. Taking cognizance of the fact that the concentration at which a line will just be visible depends upon a number of other factors besides the nature of the element, another means of standardization was adopted as follows: A graded series of standard samples was prepared with known amounts of impurity in the same base material as the alloy to be analyzed. This series of standards was sparked under the same conditions as the alloy in question and all spectra recorded on the same plate. This practically eliminated all variables of development and plate sensitivity and reduced the variables of excitation materially. The impurity content was determined by a visual comparison of the spectra of samples and standards.

The method of Meggers, Kiess and Stimson, known as the comparison standard method, is still used extensively. It is subject, however, to limitations as to reliability, sensitivity, and precision. It is not always possible to make up homogeneous alloys for standards, and since only a minute amount of sample is excited by the spark it may be difficult to excite a representative sample. This can be obviated by using solutions instead of alloys. Spark excitation is limited in its sensitivity and it is frequently necessary to analyze for smaller amounts than can be detected in the spark spectrum. The use of arc excitation usually overcomes this difficulty and furthermore the arc is somewhat more suitable for work with solutions, especially with small volumes. Visual comparison of sample with standard is limited in precision. In recent years, however, the use of the densitometer (*vide infra*) or of the logarithmic sector has increased the precision of comparison so that it is possible to make spectrochemical analyses with a precision of better

---

[1] Meggers, Kiess and Stimson, *U. S. Bureau of Standards Paper* No. 444 (1922).

than ± 5 per cent of the amount determined. In some cases a precision of ± 1–2 per cent has been claimed.

Having traced briefly the development of spectrochemical analysis, we shall now describe the functions of a modern spectrochemical laboratory and illustrate these functions by examples from the experience of the Bell Telephone Laboratories. At the outset it should be pointed out that this laboratory is confronted with an exceptionally wide variety of samples to be analyzed, due to the large number of materials used in the Bell System. Other spectrochemical laboratories, generally speaking, place a different emphasis on the various functions to be described. Furthermore, we work in cooperation with a resourceful and skillful analytical laboratory and a well organized microchemical laboratory,[2] both of which have developed methods for some materials covering the usual range of the spectrograph. As a result our functions have been limited in some directions but extended in others, so that again our experience may be somewhat different from that of other laboratories. On the other hand, the work of this laboratory illustrates the flexibility of the spectrochemical method by demonstrating that it can sometimes be even more useful in collaboration with other methods than in competition with them.

1. Qualitative Analysis for Metallic Impurities

This is probably the most common type of service performed. The spectrograph is admirably suited for such analyses since, although it utilizes only a small amount of sample, it can detect almost infinitesimal quantities of the metals and most of the metalloids. Furthermore, a complete qualitative analysis can be run in a few hours—if the spectrum is simple or the sample relatively pure one hour is usually sufficient. Thus in a relatively short time and using only a small portion of a completely unknown specimen a guide can be furnished for quantitative analysis by spectrochemical or other methods. As we shall see in the following section, it is possible also to obtain a fair idea of the quantities of each metal present from the same spectrogram.

Numerous examples could be cited but it will be sufficient to state that all "general unknowns" are first subjected to a spectrochemical qualitative examination before further analysis is attempted, and to give two typical cases. A silver contact was behaving abnormally in tarnish tests and was submitted for analysis. A qualitative test revealed that the silver contained thallium. Subsequent checks on the spools of silver wire from which contacts had been made showed

---

[2] Clarke and Hermance, "Microchemical and Special Methods of Analysis in Communication Research," *Bell Sys. Tech. Jour.*, *15*, 483 (1936).

that the wire on one spool also contained thallium and this wire was rejected for further use as contact material. Total time spent: three hours.

A lot of zinc straps for supporting aerial cables was rejected in the field because the straps were "brittle," i.e., when flexed they would crack rather than bend freely. Between areas which cracked there were areas which bent easily and showed no signs of breaking. Spectrochemical examination of the zinc at the point of fracture revealed the presence of mercury, while this element was absent in the flexible areas (Fig. 1). Time spent: two hours.

2. QUANTITATIVE ANALYSIS BY ESTIMATION

The first question usually asked when a qualitative analysis shows the presence of an element is, "How much is present?" It is easy to to tell from a qualitative plate whether the element in question is present as a major component, a minor component, an impurity, or merely as a trace. This information is frequently the deciding factor as regards further analytical work. This is particularly important in diagnosing troubles in field complaints. Very often the spectrochemical evidence alone will settle the matter—if not it will almost invariably tell just what further analyses are necessary, thus promoting the greatest possible efficiency in reaching a final solution of the problem.

The method of making such a rough quantitative analysis is to rely upon the experience of the analyst in judging the percentage of the element necessary to give the observed density of the lines under the conditions used. While such a procedure is theoretically unsound, an experienced analyst can make a surprisingly accurate "guess" by this method as has been repeatedly demonstrated by subsequent quantitative analysis.

An excellent example of an application of this technique of estimation is the identification of contact materials. It is sometimes desirable to identify the material composing a contact without removing the apparatus from its mounting, without impairing the contact for further use, and sometimes without more than momentarily interrupting its operation. The method used is to rub the surface of the contact with a small piece of fine, especially pure abrasive paper and burn the paper in a graphite arc, photographing the resulting spectrum. From the lines appearing in this spectrum which do not appear in the spectrum of the abrasive paper alone the elements composing the contact can be identified and their proportions estimated (Fig. 2).

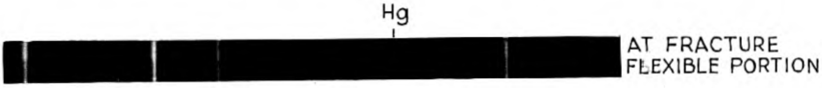


Fig. 1—Field complaint on failure of zinc straps for supporting aerial cables.

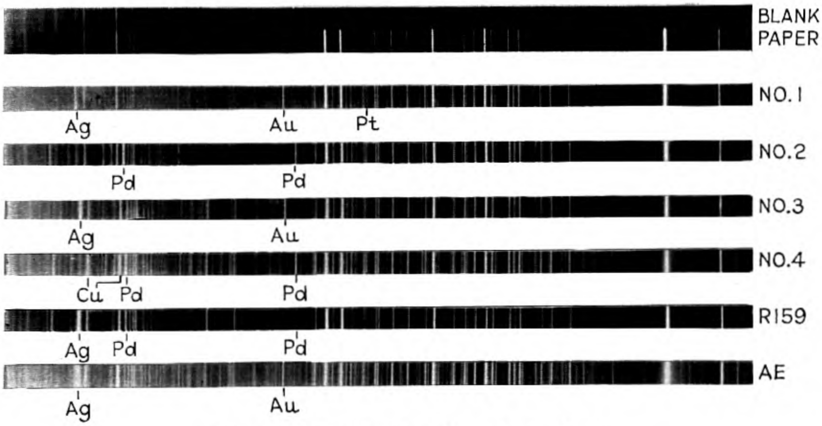


Fig. 2—Identification of contact alloys.

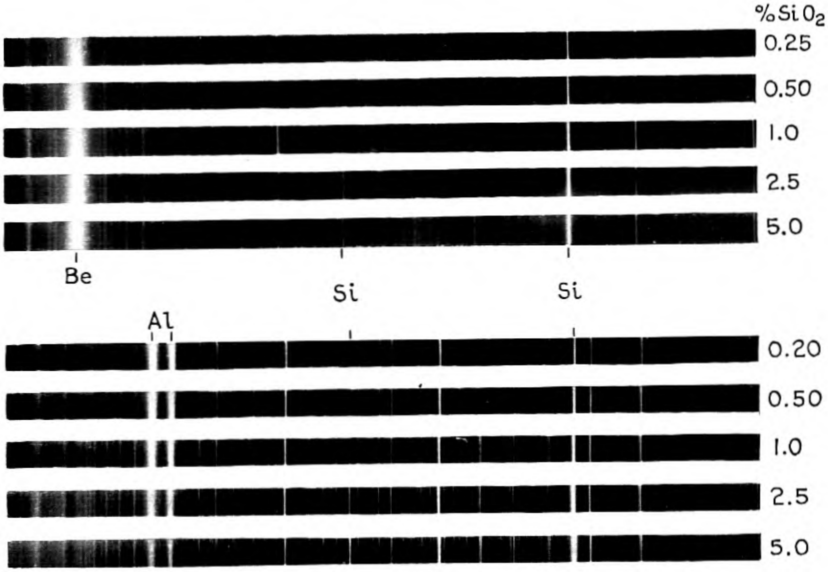
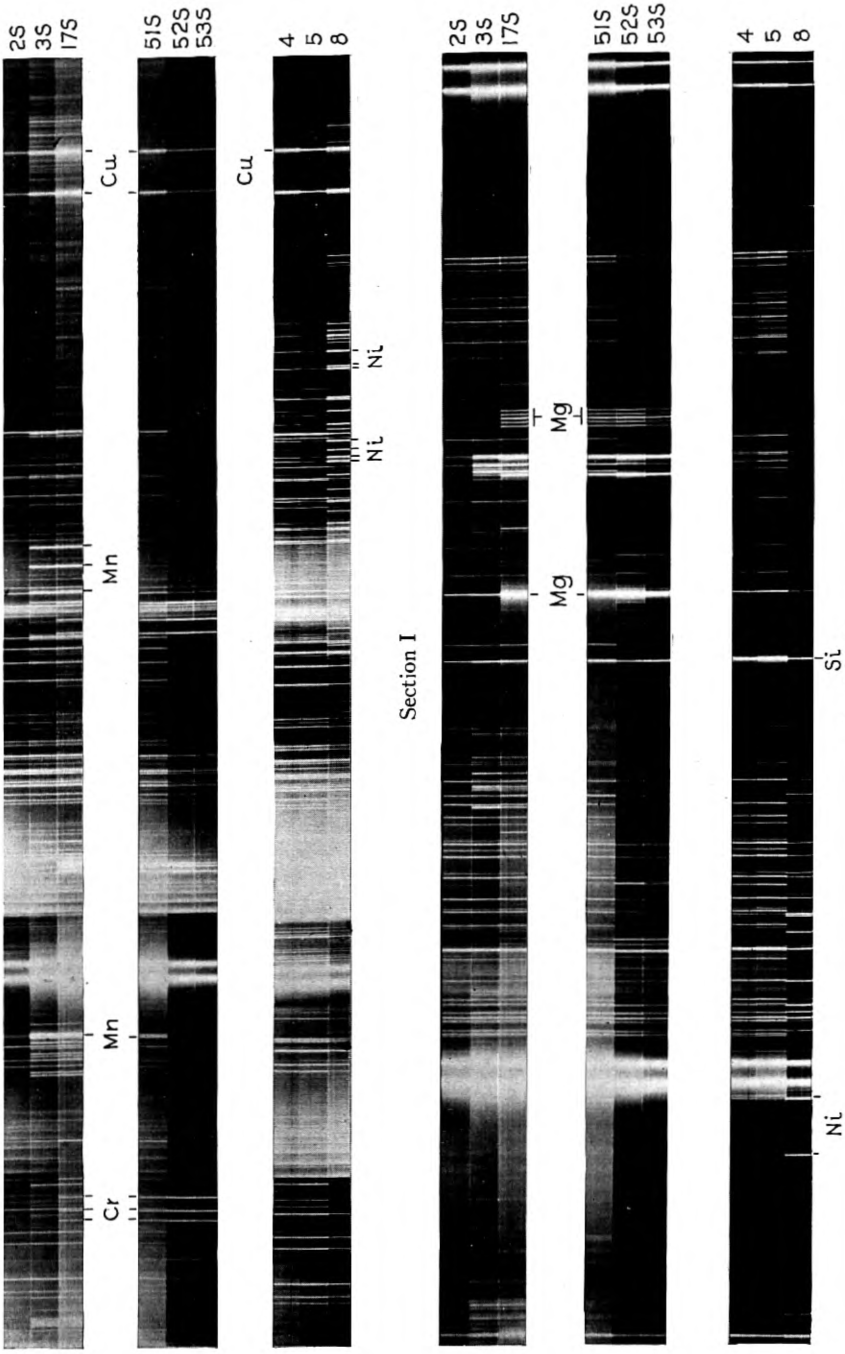


Fig. 3—Silica in beryllium oxide and aluminum oxide.



Section I

Section II

Fig. 4—Identification of aluminum alloys.

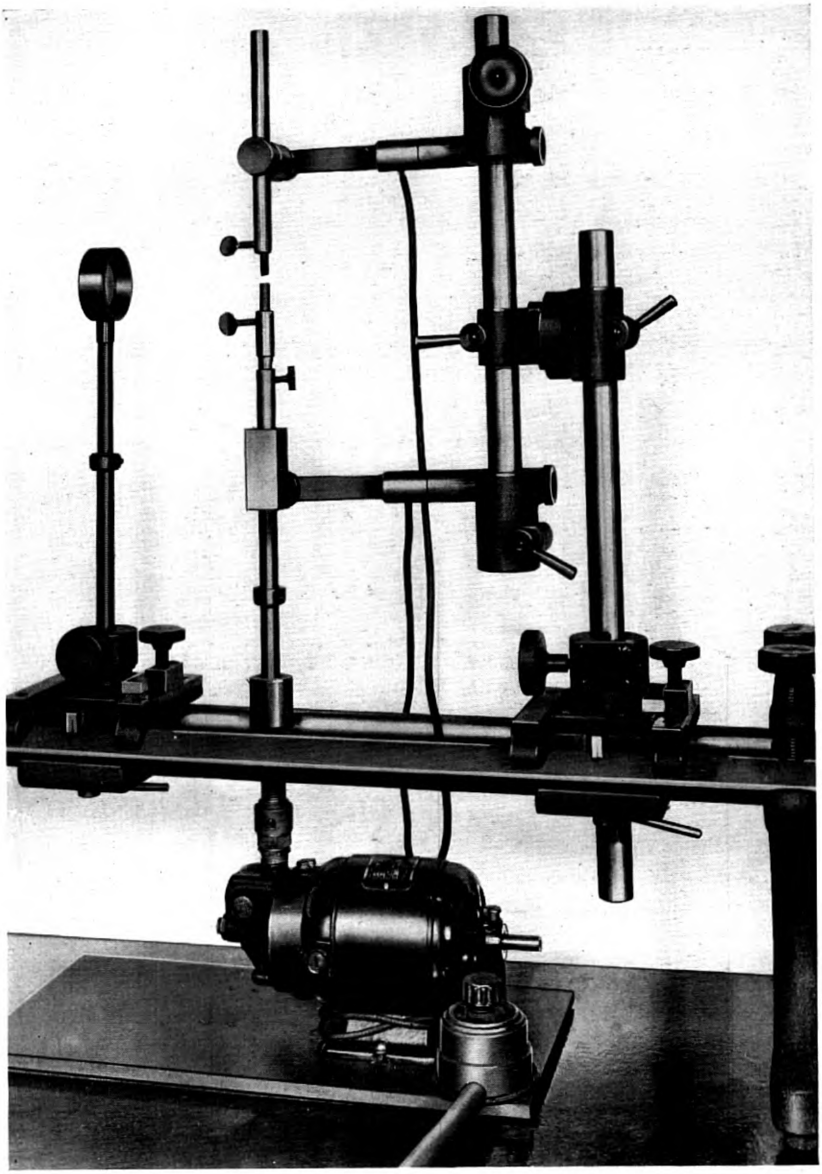


Fig. 5—Rotating electrode assembly adapted to De Gramont arc and spark stand.

Fig. 6—Step sector disk used for plate calibration.

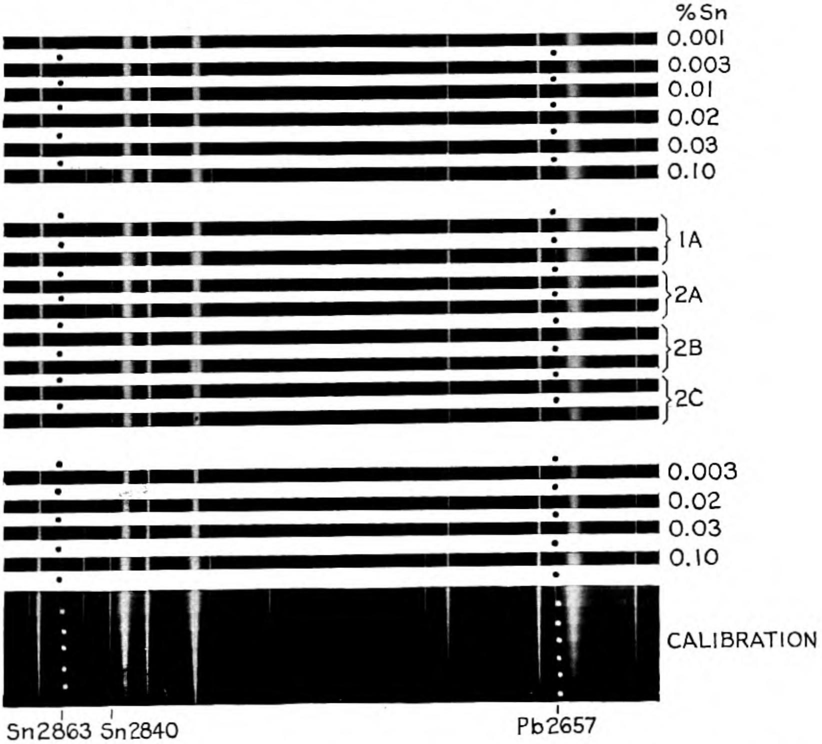


Fig. 7—Analysis of lead-antimony cable sheath for tin. An application of the internal standard method with plate calibration by the step sector.

Another example of this sort of analysis is the identification of aluminum alloys. It is a not uncommon occurrence for an alloy supposed to be of a certain composition to exhibit physical properties which would not be expected of such an alloy. Frequently it is important, therefore, to check the composition of the alloy and still preserve it for other tests. Now in the Bell System there are nine commonly used aluminum alloys, each of which contains different proportions of copper, manganese, magnesium, silicon, chromium and nickel. By examination of the spectrum obtained from a few grains of filings the analyst can easily tell by estimating the amounts of these minor components which of the nine alloys comprises the sample in question (Fig. 4).

3. Quantitative Analysis With High Speed and Moderate Precision

The greatest advantages of the spectrochemical method are speed, sensitivity, and flexibility of application. In many cases if a new material can be analyzed quickly only a moderate degree of precision is necessary. To fulfill such conditions an application of the method has been developed[3] which can be used for almost any material that can be dissolved in common solvents. The sample is first analyzed by method (1) or (2) above to find its approximate composition if that is not already known. Then a concentrated solution of the sample is prepared.

Another solution is prepared from pure specimens of the base materials present in the sample. To portions of this pure solution known amounts of the impurities present in the sample are added so that standard solutions are available for comparison with the sample. Spectra of the sample solution and the standard solutions are photographed on the same plate, using arc excitation on graphite electrodes, and the percentage of each impurity is estimated by comparison of line intensities. The method is capable of a precision of better than ± 10 per cent of the amount determined and can frequently be applied with little modification in cases where chemical methods would require considerable development work before they could be used. Furthermore the total sample needed for analysis is always very much less than is needed for a chemical analysis.

Since this is the quantitative method most widely applied in this laboratory numerous examples could be cited. To mention a few: aluminum, arsenic, tin and zinc in lead-base alloys, zinc in tin-base alloys, zinc in aluminum-base alloys, and cadmium, lead, tin, copper,

---

[3] Nitchie, C. C., *Indus. & Engg. Chem.* (*Anal. Ed.*), *1*, 1 (1929).

iron, and magnesium in zinc-base alloys have been determined by this method.

4. ROUTINE ANALYSIS

In certain cases the spectrochemical method is superior to wet methods for routine quantitative analysis. In such cases dependence is placed on a set of solid standards whose composition is known, a set of standard solutions as in the foregoing section, or a standard working curve prepared by plotting the logarithm of the relative intensity of impurity lines as compared to lines of the base metal against the logarithm of the concentration of the impurity. In the latter case recourse is had to a photometric means of determining relative line intensity. The most common methods are the logarithmic sector method and the densitometer method (see next section). Both require considerable work to prepare the standard curve, but if many samples are to be analyzed periodically a saving in time results. The precision in such cases may be better than ± 5 per cent of the amount determined.

A few cases follow in which the spectrochemical method has proved superior for routine uses: (1) the analysis of zinc-base alloys, where it is more rapid; (2) the determination of zinc in tin-base alloys and in aluminum-base alloys, where no chemical or microchemical methods are available for the smaller amounts; (3) the determination of silicon in beryllium and aluminum oxides, where it is much more rapid (Fig. 3); (4) the determination of tin in lead-antimony cable sheath where it is more rapid; and (5) the determination of magnesium in nickel alloys where it is more economical of time and sample.

5. QUANTITATIVE ANALYSIS WITH HIGH PRECISION

Of recent years much of the published research on spectrochemical analysis has been concerned with attempts to improve the precision of the method. Applications of the densitometer and of the logarithmic sector have achieved something in this direction. Both of these devices have been used with some success in this laboratory. We have also developed a rotary electrode assembly (Fig. 5) which achieves a fairly steady mean position of the arc by causing it to vibrate rapidly about a point on the optical axis rather than to wander in a random manner over the electrode surface. The apparatus is applicable to the analysis of solutions in the graphite arc and to the analysis of alloys which can themselves be used as electrodes. Salts and other loose powders of course cannot be rotated at high speeds (600 R.P.M. is recommended).

The highest precision has been claimed by Duffendack and his co-workers,[4] who use a "stepped diaphragm" to calibrate each plate with an intensity pattern in order to restrict their density measurements to the straight-line portion of the characteristic curve of the plate or to correct for deviations from the straight line portion. In this laboratory a step sector (Fig. 6) has recently been used to put the intensity pattern on each plate and the characteristic curve (Fig. 8) determined from this, using a photoelectric densitometer built in this laboratory. It is obvious from this curve that the measurement of the ratio of intensities of a given pair of lines varies with the densities of the lines as recorded on the plate. In this way errors of considerable magnitude may be introduced unless a calibration is made of the response of the plate to various levels of intensity of light of the wavelengths concerned. Results to date have shown an increased precision in all cases where plate calibration has been used, and a precision of better than ± 5 per cent of the amount determined is not uncommon. Examples from our experience include tin in lead-antimony cable sheath and other lead alloys (Fig. 7), magnesium in

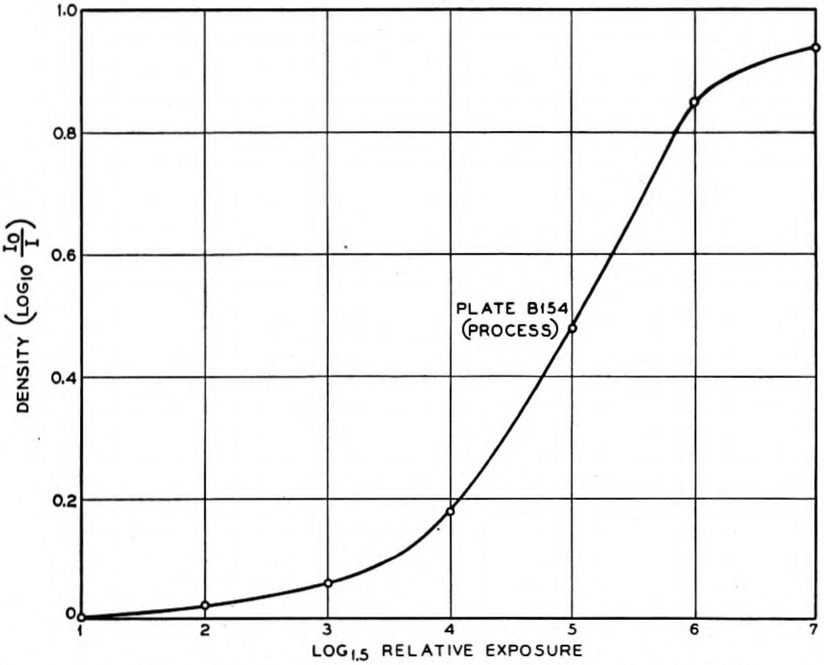


Fig. 8—A typical characteristic curve (exposure = intensity × time).

---

[4] Duffendack, Wolfe, and Smith, *Indus. & Engg. Chem.* (*Anal. Ed.*), *5*, 226 (1933) and others.

nickel filaments, and strontium-barium ratios in vacuum tube filament coatings.

6. Quantitative Analysis of Minute Traces

It has often been pointed out that spectrochemical methods can be used to assist classical and micromethods by establishing the purity of precipitates, checking the completeness of separation, demonstrating the presence or absence of interfering elements, etc. It has only infrequently been realized, however, that extremely minute concentrations of impurities can be determined by a combination of the two methods. In many cases a chemical separation can be made in which all of the impurities are removed from the bulk of the base material (it is not essential that *all* of the base material be separated as it is in chemical analyses) and can then be concentrated to fall within the spectrochemical range. In this way impurities can be determined by, say, method (3) above which could not be detected by direct excitation of the original sample. Scrupulous cleanliness and the carrying through of a reagents blank are, of course, essential when dealing with such small quantities.

7. As a Research Tool

Many times the spectrochemical method can be applied as a probe in obtaining important clues in a research project. This is particularly true when two presumably identical materials show unexpected differences in behavior. The research man may use nominally pure materials which are never really pure and thus introduce unsuspected impurities into his system which vary with different lots of materials. The spectrograph readily shows up differences in composition of the final product and thus often indicates the beneficial or detrimental effect of impurities.

In certain types of research on thin layers, the spectrochemical method is the only one of sufficient sensitivity to be used for quantitative analysis of the layers. Thus in thermionics it has been possible to measure the amounts of material on tungsten corresponding to changes in work function of the surface. Many other applications to thin films are possible and some are being worked on in this laboratory.

Thus it will be seen that chemical analysis by the emission spectrum is a powerful and versatile tool which enjoys and deserves an increasingly important role in communication research.

Acknowledgment

We are indebted to Mr. E. K. Jaycox for some of the method development as well as several of the spectrograms reproduced here.

High Speed Motion Picture Photography

By W. HERRIOTT

A motion picture camera used in taking 4000 pictures per second is described. Applications of high speed motion picture photography to a variety of problems associated with the design of telephone apparatus are given. The resulting pictures in "slow motion" permit convenient and accurate analysis of space-time relationships of mechanical parts in motion otherwise too rapid to be perceived because of their transient nature. This work has related to the development of relays, switches, clutches, ringers, dials, coin collector mechanisms, contact conditions, materials testing, etc. It has been applied to studies of noise reduction in mechanisms and to research problems associated with the production of speech by the vocal cords.

IN 1874 the French astronomer Jenssen pioneered in the use of motion picture photography as a visual aid to the study of a scientific problem. Considerable uncertainty then existed in the value of the earth's distance from the sun and Jenssen employed a camera capable of taking 48 pictures in 70 seconds during the transit of the planet Venus across the sun's disc. He hoped that errors of observation would be substantially less than those inherent in visual observations of this rare and important phenomenon but the results were disappointing due to certain photographic difficulties characteristic of the materials at his disposal. He recognized, however, the value of a series of rapidly taken photographs in making evident to the eye changes in the appearance and position of objects which would not otherwise be perceived because of their transient nature.

E. Muybridge was able to demonstrate in 1878 high speed motion pictures of animals in movement. Jenssen was content to examine his pictures singly under the microscope using the individual photographic images only for record purposes. Muybridge, however, made use of a simple viewing device in which the different pictures secured from a battery of cameras were viewed consecutively when mounted on the inner surface of a rotating cylinder provided with viewing slits. Again we have the application of photography to the discernment of transient movement not otherwise perceptible to the unaided eye. In the following year Muybridge was able to demonstrate the projection of motion pictures onto a viewing screen.

Many workers took up these pioneering experiments and improved devices were developed representing a continuous advance in the motion picture art until today we have a large industry applying the knowledge gathered by these workers to the educational and entertainment field. Paralleling commercial development of the motion picture there has been a continuous advance in its application to scientific and engineering problems, one phase of which relates to high speed motion picture photography.

Amateur and professional motion pictures are taken and projected at the rate of 16 or 24 frames or pictures per second. If pictures are taken at the rate of 48 per second and projected onto a viewing screen at the rate of 24 per second a magnification of the time axis by a factor of two occurs. The visual impression secured will be that of the same occurrence taking twice as long. If 480 pictures are taken per second and projected at the rate of 24 per second, the time magnification is 20. It is this magnification of the time axis which characterizes the picture as a "high speed motion picture." Cameras have been developed which will, under highly specialized conditions, extend the time axis by a factor of 2000 or 3000 times, although such phenomenal speeds of taking impose serious restrictions upon the nature of problems on which they may be used.

Motion pictures are usually made in a camera of the so-called "intermittent" type which refers to an intermittent motion given to the film by the film driving mechanism. An intermittent motion is employed in order that the film may be stationary during the brief exposure portion of the operating cycle after which it is rapidly accelerated and moved to an adjoining section for the next exposure. Mechanical difficulties limit the speed of operation of intermittent film moving mechanisms to a maximum taking rate of approximately 200 pictures per second. High speed motion pictures offering a magnification of the time axis by a factor of 10 can, therefore, be secured with specially constructed intermittent cameras. Higher speeds require the abandonment of the intermittent mechanism and the use of a non-intermittent or continuous film drive mechanism together with means for securing the sharp images required.

If film is moved continuously past an exposure aperture in which lies a stationary image of an object, obviously only streaks will result. Some means must be employed either to illuminate the object brightly for a sufficiently short length of time to avoid blurring of the photographic image or some device must be incorporated in the camera which will cause the image formed by the camera lens to be sharply focused on the film and to move in the direction of film travel with the same velocity. Cameras operating on both principles have been developed and are used in scientific and engineering research. A large part of the work now being done by other workers in this field is with taking speeds extending from a few hundred pictures per second to 2,000 or 2,500 pictures per second.

Cameras of the first type, that is those in which the film is driven continuously and in which the object is brightly illuminated for a short length of time (usually of the order of 2 to 10 microseconds) are relatively simple in construction. Provision is made for the reel of unexposed film which may be 25, 50 or 100 feet in length and have either 16 millimeter or 35 millimeter width. The film may be guided past a fixed exposure aperture or around a rotating drum or toothed sprocket from which it passes to the take-up reel. Power may be supplied through an electric motor either to the take-up reel or to the toothed sprocket. If a stationary exposure aperture is used, two drive sprockets, placed one above and one below the gate, may be employed. Periodic flashing of the illuminant may be secured by the use of a commutator actuated by the camera mechanism.

The second type of camera above referred to employs an optical intermittent to produce the image movement required to avoid blurring when the film is continuously driven past the exposure aperture. A variety of optical intermittents has been developed employing either lenses, mirrors, prisms, or a plane parallel glass plate or block. Regardless of choice of means, the optical intermittent serves to produce a series of rapidly moving images which move with the film velocity. Of these available methods perhaps the simplest is the plane parallel glass plate or block which, by reason of its thickness, deviates or displaces an inclined ray in a manner nearly proportional to rotation about a chosen axis.

At Bell Telephone Laboratories a high speed camera has been developed which normally operates at a taking speed of 4,000 pictures per second. This camera employs optical compensation of the type in which a cube of glass is rotated at a high rate of speed (60,000 r.p.m. for 4,000 pictures per second) between the camera lens and the sprocket as shown in Fig. 1. The compensator cube has four polished faces, each parallel to its axis of rotation and parallel to the axis of the film sprocket. One picture is taken for each quarter revolution of the compensator. The index of refraction of the glass and the dimensions of the cube are chosen to cause correct movement of the image as the film is continuously advanced past the exposure area of the sprocket. The cube rotates in the direction of the arrow which is opposite to that of the sprocket. Downward movement of the image formed by the camera lens results from change in refraction of light rays at opposite faces of the cube as the cube rotates. When the cube is in the position shown at "*A*" an upward displacement of the image to the point "*a*" results. Rotation of the cube in the clockwise direction diminishes the amount of displacement with the result that a downward image movement takes place which is synchronized with the film movement. With the compensator cube in position as shown at "*B*," that is, its entrant and emergent faces perpendicular to the optical axis, no vertical deviation or displacement of the image results and the image falls at the point "*b*" on the film sprocket. Further rotation of the cube causes the image to move downward to the point "*c*" where the exposure is terminated and the next adjacent face of the cube comes into play. In this manner a succession of images is laid down frame by frame at a high rate of speed, each elemental area of the film having received exposure during a substantial part of the rotation cycle. The duration of each exposure is controlled by the film speed and by the angular height of a fixed aperture in front of each of the four faces of the cube. The film sprocket is directly driven from the motor shaft. Spur gears are employed to drive the optical compensator. A separate motor is employed to drive the take-up reel. 16 millimeter film in hundred foot lengths is used.

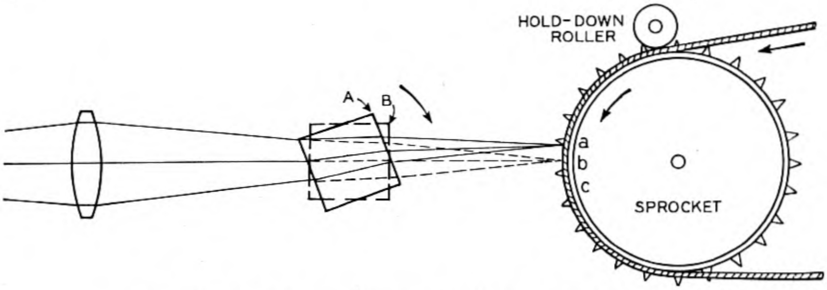


Fig. 1—Schematic arrangement of high speed camera.

A finder is provided which permits viewing the image on the film as projected upon a ground glass screen mounted on the hinged door of the camera. Lenses of various focal lengths are interchangeable on the front of the camera. The camera is mounted upon a substantial tripod and is readily portable. Figure 2 shows the exterior of the Bell Telephone Laboratories high speed camera. Figure 3 shows the interior of this camera where the location of the film spools and main drive sprocket are shown. Figure 4 shows the camera with its two motors and portable lighting units of the type developed for

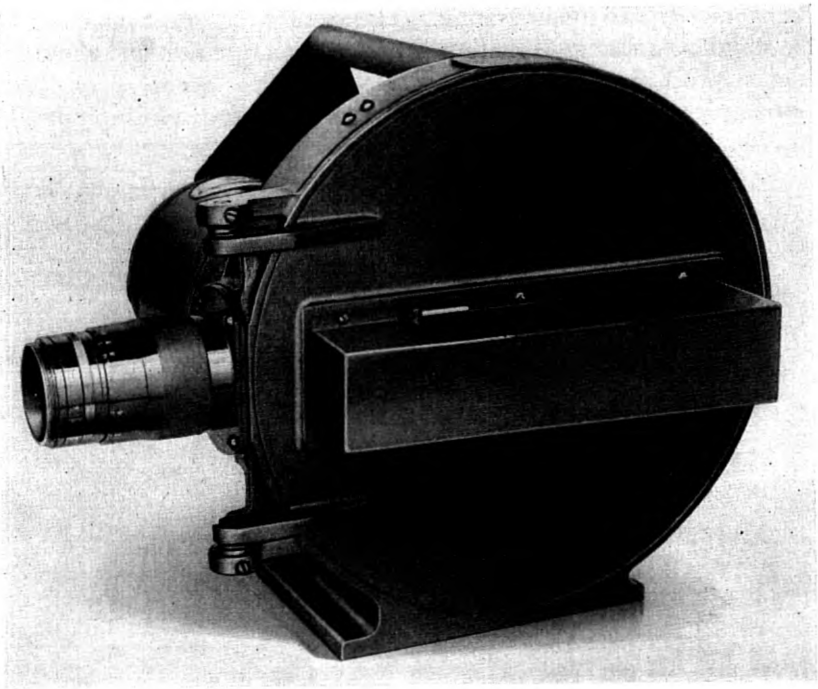


Fig. 2—High speed camera used in taking 4000 pictures per second.

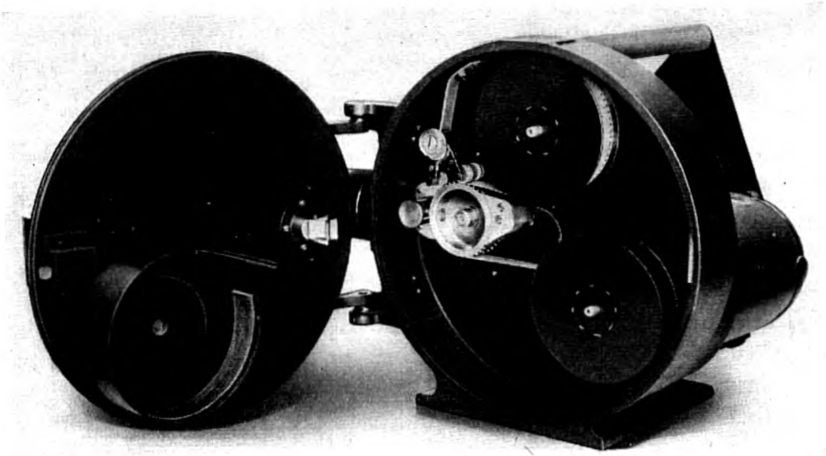


Fig. 3—Interior view of camera.

use in high speed photography. The camera has been adapted to use standard amateur 16 millimeter motion picture film and the commercial processing service of the film supplier is used. High-speed motion pictures in color are readily made, utilizing available films and processing.

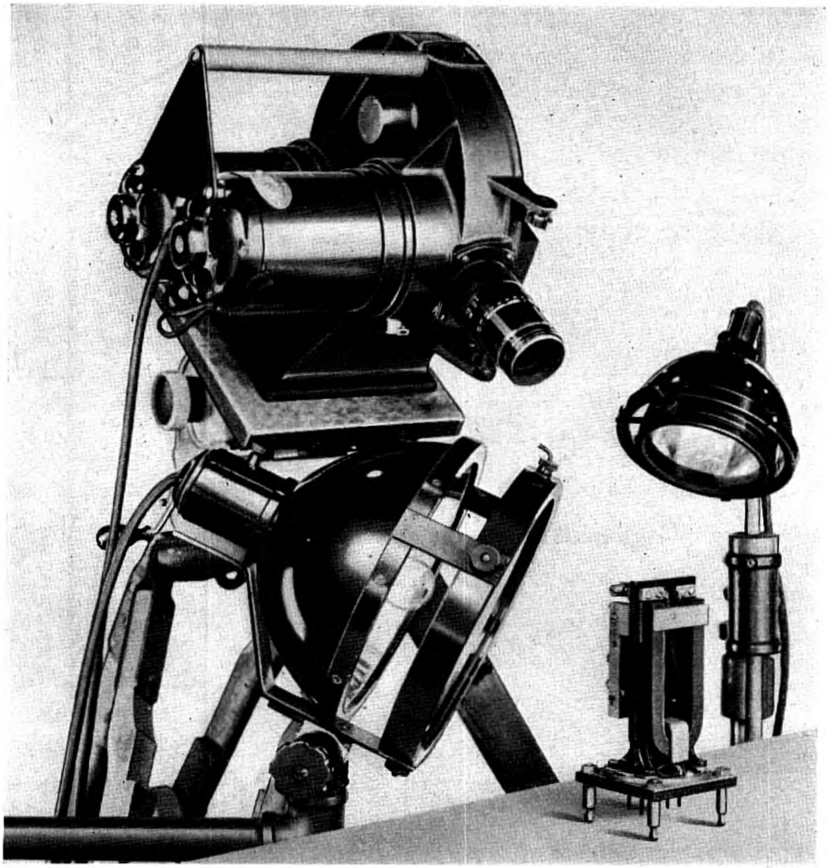


Fig. 4—Camera with portable lighting units used for high speed work.

The optical intermittent type of camera has certain advantages over other types. The unit is self-contained and is independent of lighting equipment. It can be made relatively light in weight and highly portable. Further, it can be applied to the study of self-luminous sources as required in the study of arc lamps, lamp burnouts, fuse burnouts and other problems where the phenomena under investigation are self-luminous.

High speed motion picture photography is finding extensive application in industry where it is applied to "motion analysis" of a variety of manufacturing operations and to problems associated with the design and performance of machinery. A particular value of this method of motion analysis lies in the convenience and accuracy with which space-time relationships of moving parts can be determined. Individual frames can be enlarged as photographic paper prints or projected onto a lined screen; in either case, the movement of parts can be measured with a high degree of accuracy. Extensive application is made in automotive and aeronautical engineering to the study of problems associated with fuel combustion, vibration in motors, airflow around structures and to propeller design and performance. It is coming into use in the fields of biology and medicine, especially in the study of muscular and nervous reactions. It has been applied to microphotography in the photographing of biological specimens at magnifications of 500 to 700 times at a taking rate of 1000 pictures per second.

In high speed photography one of the principal problems has been the securing of adequate illumination of the subject. Exposure times are of the order of one twenty-thousandth second or less for taking speeds of 4000 frames per second. Artificial illumination must be used and arc lamps and tungsten incandescent lamps are employed. Special projection bulbs are secured which are greatly over-voltaged during the few moments required in the taking of a picture. The lamps are designed for short life and to operate as close to the melting point of tungsten as is practicable. Two lamps are operated in series during setting up and adjustment of the apparatus and full voltage applied to both lamps at the moment of taking by the use of series-parallel switching arrangements. The lamps are housed in specially designed lighting units employing high aperture lenses or mirrors which serve to image the source directly on the object at a desired magnification. Provision is made for the reduction of heat by the use of water cells of suitable thickness. Excessive heat in the image will frequently cause distortion in delicate apparatus which must be avoided. One or more lighting units may be employed depending upon the size of the object being photographed, the taking speed, the lens aperture and the film speed. Brightnesses of the order of 10,000 to several hundred thousand foot candles are frequently employed.

At Bell Telephone Laboratories, high speed photography has been applied to a wide variety of problems associated with design and performance of telephone apparatus. Pictures have been taken of standard equipment and of experimental equipment in course of development. Much valuable data not otherwise obtainable have been secured regarding functioning of telephone equipment under use conditions. This work has related to relays, switches, clutches, ringers, dials, coin collector mechanisms, contact conditions, etc. It has been applied to the testing of materials in connection with impact testing, stress analysis and bending moment. It has also been applied to reduction of noise in apparatus, particularly in machines of the type widely employed for accounting purposes, coin counting, typewriter mechanisms, and other high speed operating mechanisms where it was desired to analyze and remedy certain noise conditions. It has found particular use in fundamental research associated with the production of speech by the vocal organs. High speed motion pictures have been made of the vocal cords in action employing taking speeds of 4000 frames per second. These pictures are particularly valuable to teachers of speech and music and to the medical profession. Such pictures have, of course, great value to engineers working on problems associated with the transmission of speech over telephone circuits.

A service in high speed photography is available to engineers of the Laboratories as a visual aid to their study of problems associated with the design, manufacture and performance of telephone apparatus. The high degree of portability which has been achieved in both the camera and lighting equipment lends itself well to extensive application of this service to engineers. Figures A to L show series of individual frames illustrating a variety of problems to which this service has been applied.

High speed motion pictures of a variety of subjects have been taken in color. Color pictures have been made of vocal cords and of photoelastic effects revealed in transparent materials under polarized light. Such pictures are valuable in studying stress and impact conditions as affecting design of equipment. Stereoscopic high speed motion pictures have been made both in color and in black and white.

High speed motion picture photography is finding increased application to a variety of problems associated with the work of the Laboratories. It is believed that more extended applications of its use will follow.

TYPICAL HIGH SPEED MOTION PICTURES

A—High speed photographs of an experimental model of the step-by-step switch showing the wiper rising to the cut-in position and overthrowing on the first contact of its associated bank. Motion pictures of this type disclosed in great detail movements of the wiper involved in the operation of this switch. These movements are so rapid that but little is learned from a visual examination of the switch while in operation. 4000 pictures per second.

B—These pictures show normal action of the vertical pawl in the step-by-step switch. Again, high speed motion pictures show much that cannot be gained from visual examination. In this case, the pictures revealed whether or not the action of the pawl and ratchet was satisfactory under extreme operating conditions. 2000 pictures per second.

C—To study the conditions under which handset breakage might occur, apparatus was developed in which breakage takes place under controlled conditions. High speed photographs revealed distortions in the handset which under certain test conditions resulted in breakage at the moment of impact. Information relating to the structural strength of the handset was secured from measurements made on many individual pictures of this series. The pictures shown illustrate breakage of an experimental 3-piece E type handset after falling from a height of 5 feet and striking a rigid steel bar shown at the bottom of each picture. 4000 pictures per second.

D—Information gained from a study of high speed pictures of the type represented by C resulted in strengthening of the reenforcement of the handle at certain points which reduced the possibility of breakage under normal use conditions. 4000 pictures per second.

E—High speed photography is applied extensively to the study of explosion of gases in motors, to ballistic problems associated with the explosion of gun powder and to other rapid phenomena of a self luminous type. This series of pictures shows the melting and burning of fuse wire under heavy current conditions. 20-ampere fuse wire is shown during burn-out on direct short. The violence and extent of the action are well shown in these pictures. 4000 pictures per second.

F—Certain normally isotropic transparent materials become birefringent when examined in a stressed condition under polarized light. Extended use of this effect is made in the study of stress distributions in engineering structures and in models of mechanical parts. High speed photography is now applied to these photoelastic effects exhibited in a glyptol sample under impact stress condition. This series of pictures shows impact testing of an unnotched glyptol specimen in plane polarized light. 300 pictures per second.

G—Poor contact conditions in relays may give rise to improper circuit operation. High speed motion pictures have been useful in the study of contact chatter in relays and other similar devices. This series of pictures shows normal operation of contacts. 2000 pictures per second.

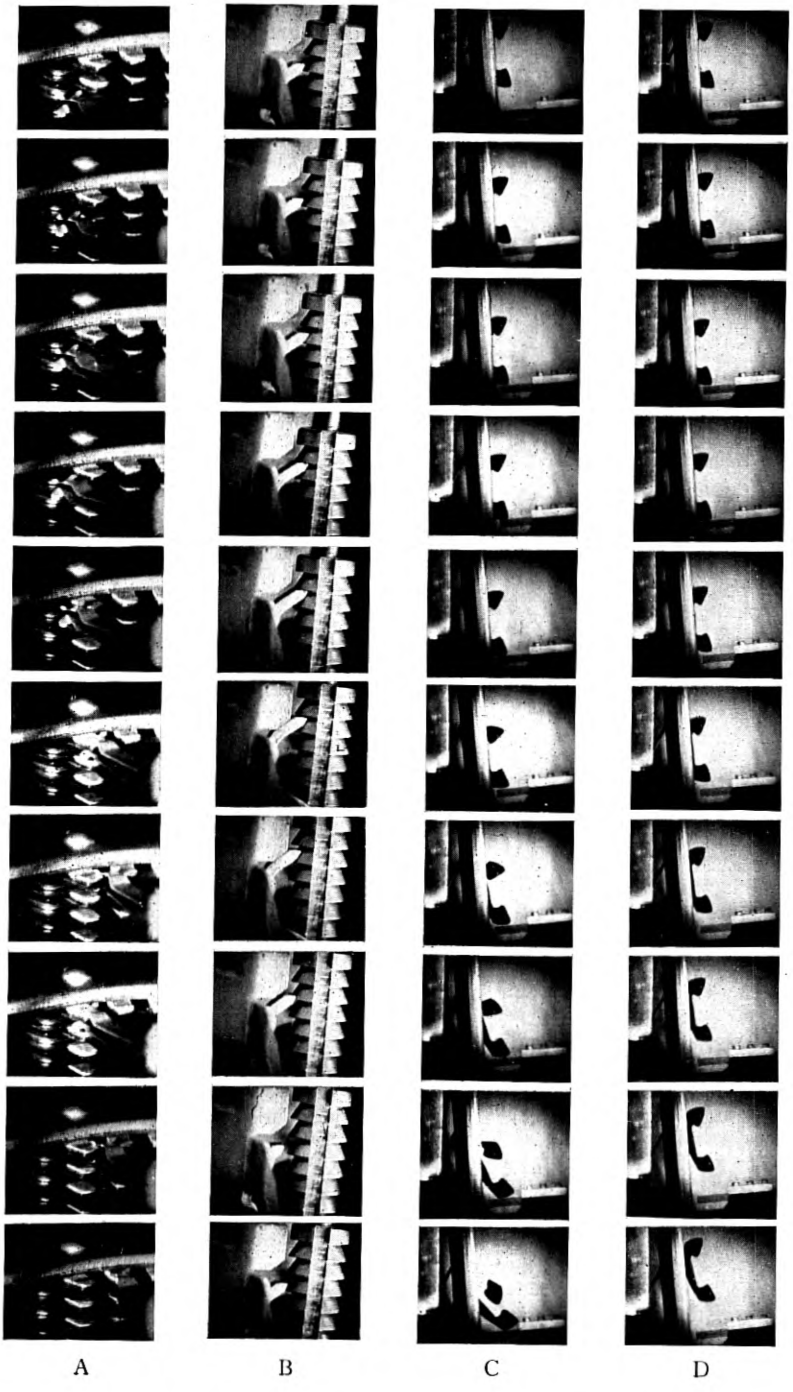
H—This series of pictures exhibits contact chatter. In the first picture of this series the movable contact spring is shown contacting the left fixed contacts. In the second and third pictures, the movable contact spring has been drawn against the right hand fixed contacts. At this point the current is cut off and the movable contact springs return to normal as shown in the sixth picture. Chatter occurs at this point with the movable spring returning to make contact with the stationary contacts shown at the right. Two cycles of chatter condition are shown. 2000 pictures per second.

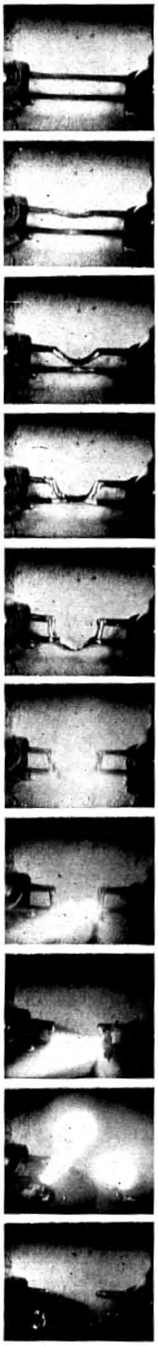
I—This series of pictures shows the No. 14 teletypewriter locking arm lever and cam during overthrow which gives rise to noisy operation. They illustrate a typical source of objectionable noise in apparatus of this type. Excessive clearance between the cam and the locking arm lever is shown which results in impact noise on the return of the locking arm lever. 1800 pictures per second.

J—This series of pictures shows a modified No. 14 teletypewriter locking arm lever and cam in which the overthrow has been eliminated with subsequent reduction in noise. It can be seen that the lever arm now closely follows the contour of the cam. 1800 pictures per second.

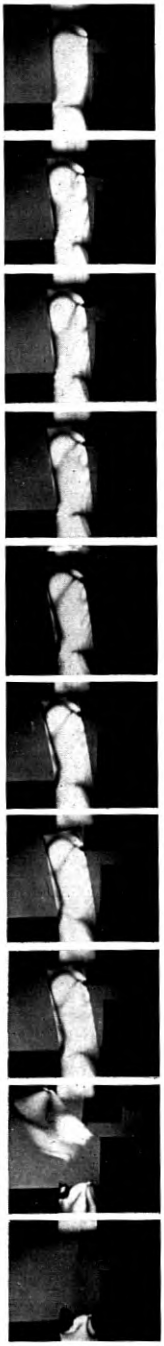
K—A knowledge of the fundamentals of speech and hearing is important to designers of telephone apparatus. High speed motion picture photography has been applied to problems associated with the production of speech by the vocal mechanism. The pictures show vocal cords vibrating in production of speech sound at a frequency of 120 cycles per second. Pictures of this type offer a unique and practical means of securing much useful information relating to the production of speech. 4000 pictures per second.

L—At L is shown the action of the clapper striking one gong of an experimental 20-cycle ringer. This picture revealed more strokes of the clapper per second of operation than was desired. This condition resulted in a peculiar acoustic effect, readily explained from this series of pictures. 2000 pictures per second.

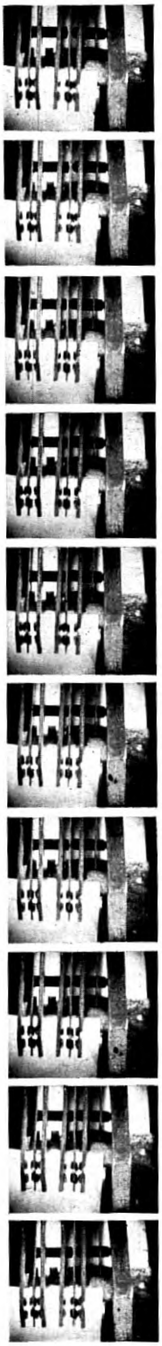

A
B
C
D



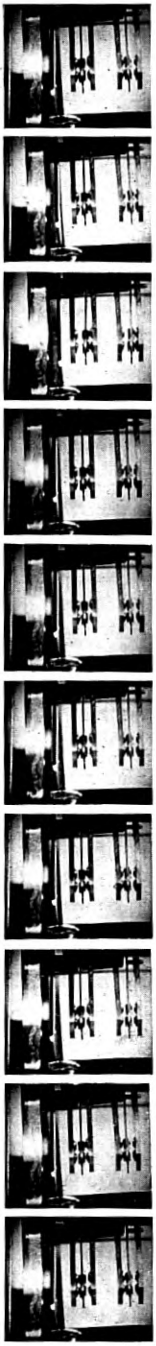
E



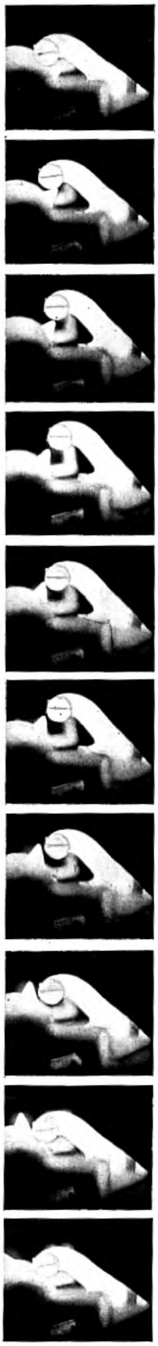
F



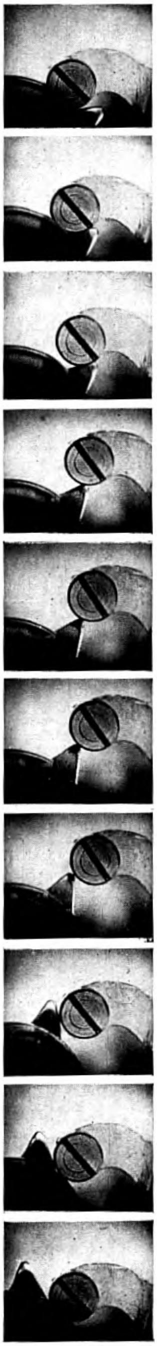
G



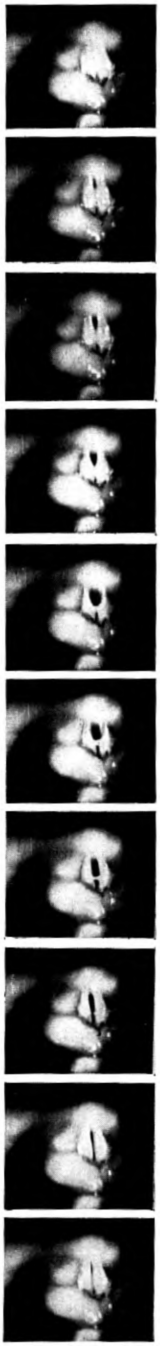
H



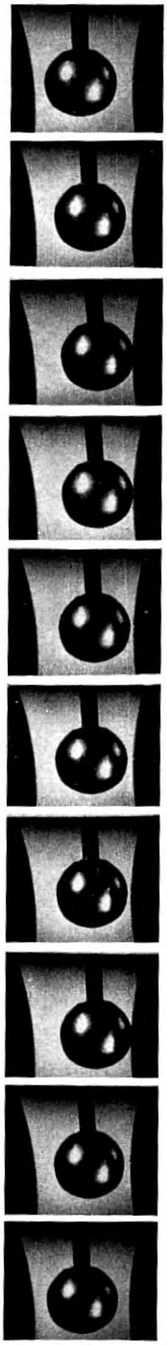
I



J



K



L

BIBLIOGRAPHY

1. "Motion Picture Camera Taking 3200 Pictures per Second," C. F. Jenkins, *Trans. Soc. Motion Picture Engineers*, No. 17, Oct. 1923, p. 77.
2. "The Chronoteine Camera," C. F. Jenkins, *Jour. S. A. E.*, No. 22, Feb. 1928, p. 200–2; *Trans. Soc. Motion Picture Engineers*, No. 25, Sept. 1926, p. 25.
3. "New High Speed Kinematographic Camera," T. Suhara, *Proc. Imp. Acad.*, Tokio, No. 5, Oct. 1929, p. 334–7.
4. "New Ultra-Speed Kinematographic Camera," T. Suhara, *Report No. 60, Tokio Univ. Aeronautical Res. Inst.*, May 1930, p. 187–94.
5. "A Non-intermittent High-Speed 16 M.M. Camera," F. E. Tuttle, *Jour. Soc. Motion Picture Engineers*, No. 21, Dec. 1933, p. 474.
6. "The N. A. C. A. Photographic Apparatus for Studying Fuel Spray From Oil Engine Injection Valves and Test Results From Several Researches," E. G. Beardsley, *N. A. C. A. Tech. Report No. 247*, 1927.
7. "The N. A. C. A. Apparatus For Studying The Formation and Combustion of Fuel Sprays and the Results From Preliminary Tests," A. M. Rothrock, *N. A. C. A. Tech. Report No. 429*, 1932.
8. "Stroboscopic and Slow Motion Pictures by Means of Intermittent Light," H. E. Edgerton, *Jour. Soc. Motion Picture Engineers*, No. 3, 1932, p. 356.
9. "The Mercury Arc as an Actinic Stroboscopic Light Source," H. E. Edgerton and K. J. Germeshausen, *Rev. of Scientific Instruments*, Oct. 1932, p. 535.
10. "High Speed Motion Pictures," H. E. Edgerton, *Electrical Engineering*, Feb. 1935, Vol. 54, p. 149–153.
11. "Precision Timing of Athletic and other Sporting Events," C. H. Fetter and H. M. Stoller, *Electrical Engineering*, June 1933, Vol. 52. p. 386–391.
12. "Time Microscope for High Speed Machines," J. P. Maxfield, *Modern Packaging*, July 1936, Vol. 9, No. 11.
13. "High Speed Motion Pictures of the Vocal Cords," W. Herriott and D. W. Farnsworth, *Radiography and Clinical Photography*, June 1938, Vol 14, No. 3, pp. 26–27.

An Optical Harmonic Analyzer *

By H. C. MONTGOMERY

An instrument which makes a Fourier Series Analysis of a function by optical means has recently been completed. The function to be analyzed is supplied in the form of a variation in the density or in the width of the transparent portion of a photographic film. The analysis is performed by a direct evaluation of the integrals which form the coefficients in a Fourier Series, and the results are theoretically exact in the sense that the measurement of each harmonic is independent of the other harmonics which may be present in the function. The operation of the instrument is largely automatic, and is rapid enough so that 30 harmonics can be measured in about a minute and a half.[1]

A PERIODIC function can be represented for all values of the variable by a Fourier Series. A function which is not periodic can be so represented between any finite limits, although the series may be entirely unlike the function beyond these limits. If a function is approximately periodic, the Fourier Series representing adjacent portions of it will generally be approximately alike.

Although in general an infinite number of terms is required to represent a function exactly, it is common experience that a great many functions of practical interest can be closely approximated by a series of from ten to thirty terms.[2]

PRINCIPLE OF OPERATION

The principle of this analyzer was suggested by E. C. Wente of these Laboratories.[3] It may be outlined as follows.

The Fourier Series expansion of a function is given by either of the following equivalent expressions.[4]

---

* Presented at Meeting of Acoustical Society of America, Washington, D. C., May 3, 1938.

[1] For comparison, analysis to 30 harmonics on the Henrici type instrument requires five or six hours. A resonance analyzer, such as the vibrating reed type, can complete an analysis in a few seconds, but the phases will not be given, and if the function is provided in graphical form it must be converted into an electrical or acoustic wave form repeated enough times for the resonant elements to reach a steady state response.

[2] A description of a number of the more important methods of harmonic analysis, together with a bibliography, is contained in "Sound Analysis," H. H. Hall, *Jour. Acous. Soc. Amer.*, vol. 8, pp. 257–262, April 1937.

[3] U. S. Patent No. 2,098,326.

[4] The expressions in this form apply when the fundamental period is 2π. There is no loss of generality, as the scale of abscissa can always be so chosen as to conform

$$f(x) = a_0 + \sum_1^\infty a_n \cos nx + \sum_1^\infty b_n \sin nx \tag{1}$$

$$= c_0 + \sum_1^\infty c_n \cos (nx - \phi_n). \tag{2}$$

Comparison of (1) and (2) gives the following relations between the coefficients in the two forms of the expression:

$$a_n = c_n \cos \phi_n, \qquad b_n = c_n \sin \phi_n. \tag{3}$$

$$c_n^2 = a_n^2 + b_n^2, \qquad \phi_n = \tan^{-1} \frac{b_n}{a_n}. \tag{4}$$

The form (2) giving amplitude and phase angle of the harmonics is generally more useful, but most methods of analysis give the coefficients in form (1) necessitating the computation of the amplitude and phase angle from the relations (4). One of the advantages of the optical analyzer is that it will give either set of coefficients directly.

The coefficients in the Fourier Series can be determined from the following expressions:[5]

$$a_n = \frac{1}{\pi}\int_0^{2\pi} f(x) \cos nx\, dx, \qquad b_n = \frac{1}{\pi}\int_0^{2\pi} f(x) \sin nx\, dx. \tag{5}$$

$$c_n = \frac{1}{\pi}\int_0^{2\pi} f(x) \cos (nx - \phi_n) dx, \quad \int_0^{2\pi} f(x) \sin (nx - \phi_n) dx = 0. \tag{6}$$

$$a_0 = c_0 = \frac{1}{2\pi}\int_0^{2\pi} f(x) dx. \tag{7}$$

We will now describe two methods by which a function can be represented on a photographic film. In a *variable area record* the film, which is elsewhere opaque, contains a transparent portion whose width at any point is proportional to the function. Such a record is shown in the upper part of Fig. 1. In a *variable density record* the

---

to this requirement. In fact this selection of a proper scale corresponds directly to a necessary adjustment of the analyzer.

[5] The expressions given in (6) can be derived from the more familiar expressions (5) as follows. From (3)

$$a_n \cos \phi_n + b_n \sin \phi_n = c_n(\cos^2 \phi_n + \sin^2 \phi_n) = c_n,$$
$$b_n \cos \phi_n - a_n \sin \phi_n = 0.$$

Substituting values of a_n and b_n given by (5) in these expressions leads at once to the expressions (6).

function is represented by gradations in the density of the film such that the light transmission at any point is proportional to the function, the density being uniform in a direction perpendicular to the axis. Such a record is shown schematically in the lower part of Fig. 1. With either type of representation of a function $g(x)$, it will be seen that the amount of light transmitted through a narrow vertical strip of width dx is proportional to $g(x)dx$. If two or more such records are superimposed, the light transmitted through all of them will be proportional to the product of the recorded functions, provided not more than one of the records is of the variable area type.

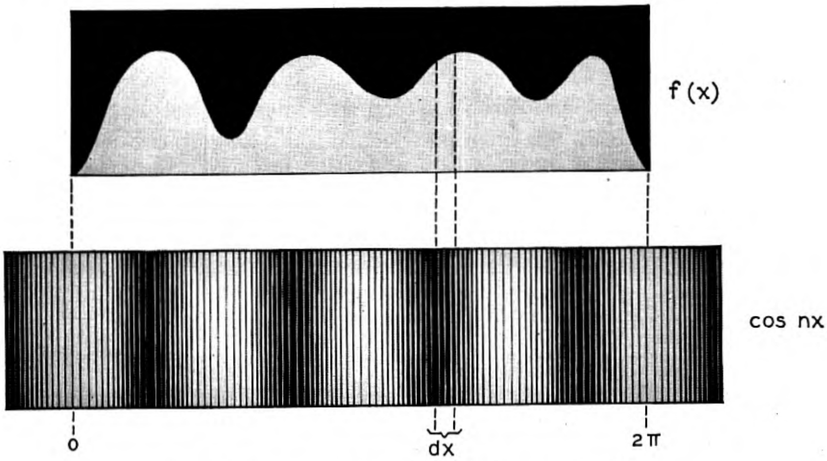


Fig. 1—Representation of $f(x)$ and cos nx on film.

The determination of a_n and b_n is now very straight-forward. Suppose we have $f(x)$ recorded on one film and cos nx on another. For illustration we will assume that $f(x)$ is recorded in variable area and cos nx in variable density, as shown in Fig. 1, although the only necessary requirement is that they shall not both be variable area. If the two films are superimposed, the amount of light transmitted through both of them between the limits zero and 2π is just the first integral in (5) and hence proportional to a_n. Similarly, if the cosine screen is moved a quarter wave-length along the axis it becomes sin nx and we have at once b_n. If the cosine screen is moved a whole wave-length along the axis, the transmitted light will go through a maximum. It will be shown below that this maximum value is proportional to c_n, and the position of the cosine screen at which it occurs is ϕ_n.

One more matter needs to be considered before we write down the expressions which describe the operation of the analyzer. Since $f(x)$

and cos nx will in general have both negative and positive values they cannot be directly represented by the transmission of light, which is essentially positive. However, the addition of a constant to each function will eliminate this difficulty, and merely results in a constant in the measured amplitude, as shown below.

The optical transmission of the film on which $f(x)$ is recorded may be written

$$A + f(x),$$

where A is a constant large enough to make the expression positive for all values of x. Similarly, the transmission of the cosine screens may be written

$$B_n[1 + M_n \cos (nx - \theta)],$$

where θ is a parameter denoting the position of the cosine screen along the x-axis, M_n is a constant somewhat less than unity, known as the modulation of the record, and B_n is a constant which is seen to be the average optical transmission of the screen.

If one or both of these records is of the variable density type, the total transmission when they are superimposed will be

$$\begin{aligned} T &= \int_0^{2\pi} B_n[A + f(x)][1 + M_n \cos (nx - \theta)]dx \\ &= \int_0^{2\pi} AB_n dx + \int_0^{2\pi} B_n f(x) dx + \int_0^{2\pi} AB_n M_n \cos (nx - \theta) dx \\ &\qquad + \int_0^{2\pi} B_n M_n f(x) \cos [(nx - \phi_n) - (\theta - \phi_n)]dx, \\ &= 2\pi B_n(A + c_0) + \pi B_n M_n c_n \cos (\theta - \phi_n). \end{aligned} \tag{8}$$

To obtain a_n, we take the difference in T for $\theta = 0$ and $\theta = \pi$, which is seen to be

$$2\pi B_n M_n c_n \cos \phi_n = 2\pi B_n M_n a_n. \tag{9a}$$

Similarly, to obtain b_n, we make $\theta = \frac{\pi}{2}$ and $\theta = \frac{3\pi}{2}$, giving for the difference in T

$$2\pi B_n M_n c_n \sin \phi_n = 2\pi B_n M_n b_n. \tag{9b}$$

To obtain c_n and ϕ_n note that the maximum value of T occurs at $\theta = \phi_n$ and the minimum at $\theta = \phi_n + \pi$, which serves to determine ϕ_n. The difference between the maximum and minimum values of T is

$$2\pi B_n M_n c_n. \tag{10}$$

If the factors B_n and M_n can be made approximately constant for all the screens, the coefficients for either form of the Fourier Series are directly proportional to the change in the amount of transmitted light for specified pairs of positions of the cosine screens.

DESCRIPTION OF THE INSTRUMENT

The process which the analyzer is required to carry out consists of superimposing the function to be analyzed on a cosine screen and measuring the variation in the transmitted light when the cosine screen is moved along the x-axis. This is repeated with a different cosine screen for each harmonic which it is desired to measure.

A schematic diagram of the instrument is shown in Fig. 2. The film containing $f(x)$ is placed in a holder at A and strongly illuminated by an incandescent lamp and condensing lens. An enlarged image of $f(x)$ is formed at B on a window bounded by two knife edges .750 inch apart. Functions of different length are accommodated by adjusting the optical enlargement so that the image of the portion of $f(x)$ to be analyzed will just fill the window. The cosine screens slide in a track directly behind the window, and receive the image of $f(x)$. The transmitted light is collected by another lens and brought to a photocell.

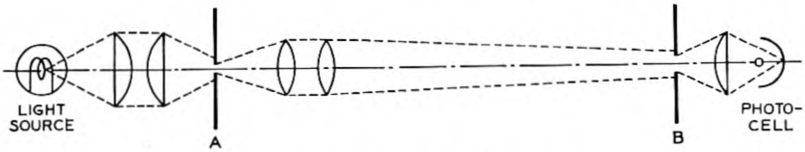


Fig. 2—Diagram of optical system.

A series of cams and levers is arranged to bring the cosine screens out of a drum shaped magazine in which they are stored into the optical path, give them the small motion required for analysis, and return them to the magazine, which is then rotated to bring the next screen into position. These operations are all automatic, and the attention of the operator is required only for the adjustment of the enlargement and focus and resetting of the cams at the beginning of each analysis. A photograph of the instrument is shown in Fig. 3.

The variations in the photocell output take place at the rate of about two cycles per second. These are recorded on a moving chart by an instrument similar to a high speed level recorder,[6] differing from it chiefly in having a linear instead of a logarithmic scale.

---

[6] "A High Speed Level Recorder," Wente, Bedell and Swartzel, *Jour. Acous. Soc. Amer.*, vol. 6, p. 121, January 1935.

The present instrument has been designed to take records of $f(x)$ which are from one-sixteenth to five-sixteenths of an inch long and no higher than their length. The focal length of the enlarging lens is 1.5 inches. The collecting lens is placed quite close to the cosine screens, and forms an image of the enlarging lens on the plate of the photocell. With this arrangement the patterns of both $f(x)$ and the cosine screens are well diffused on the photocell plate, so that surface variations in sensitivity of the plate are unimportant. The illumination is uniform across the field to ±2 per cent.

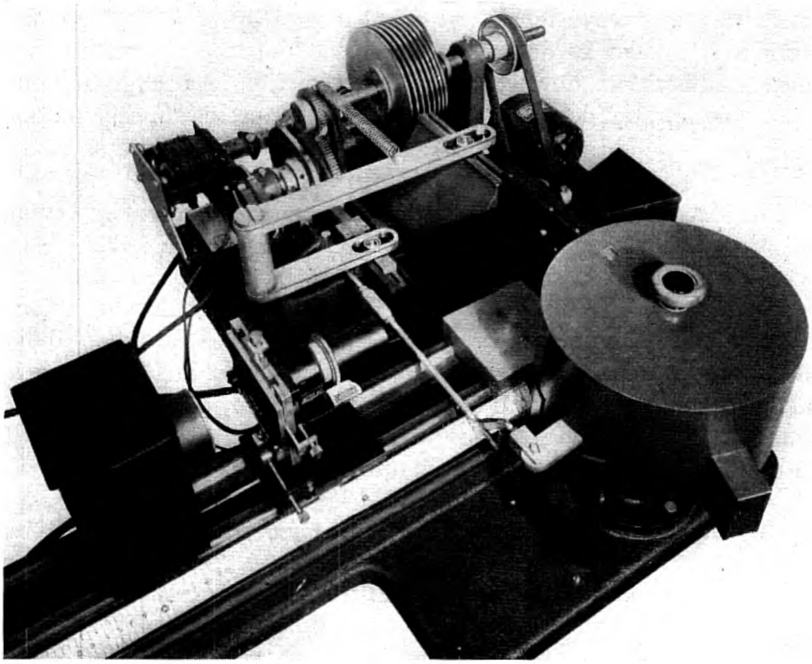


Fig. 3—The optical harmonic analyzer.

The cosine screens were made on photographic plates, by printing from variable density motion picture sound track negatives containing records of pure single frequencies. The pattern thus produced is about 1 × 2 inches. The increase in width of the track from about one-tenth of an inch in the negative to one inch on the plate was secured by making a "contact" print with negative and plate slightly separated, and moving the plate sideways under the negative while printing.

The important requirements for the screens are good wave form, uniformity in modulation and average transmission, and accuracy in wave-length. In the present instrument it was found possible to keep the harmonic content of the screens down to 5 per cent. The modulation varied from 79 per cent to 94 per cent in different screens, and the average transmission from 20 per cent to 24 per cent. Variations in the wave-length of the screens amounted to about 1 per cent.[7]

It is convenient though not necessary to have good wave form in the screens. When readings are made in pairs as described above, the effects of even order harmonics in the cosine screens cancel out. Moreover, G. R. Stibitz has shown [8] that the cosine screens may have practically any wave form (not even necessarily periodic), and correction factors can be derived for them. The process of correction is rather cumbersome, however, since the correction for each harmonic is not a constant, but depends on other harmonics present in the function.

USES OF THE ANALYZER

This instrument was designed particularly to accommodate sound records on film as used in commercial motion picture work. However, functions from any other source can be analyzed equally well if they are reproduced with the proper dimensions on film. Provision is made for measurement of the first 30 harmonics. As stated above, the function must be between 1/16 and 5/16 inch in length and no higher than its length. At the speeds customarily used for recording sound on film this corresponds to a fundamental frequency of from 65 to 310 cycles per second, or 1950 to 9300 cycles per second for the 30th harmonic.

The smallest harmonic which the instrument will indicate is about 2 per cent of the peak which it can accommodate. In connection with this statement, it should be remembered, however, that for many functions the largest harmonic in the analysis is considerably less than the peak amplitude of the function, which reduces the effective amplitude range.

An interesting check on the operation of the analyzer can be obtained by making an analysis of a simple geometric wave form. For example, a single cycle of a saw-toothed wave can easily be formed by placing a straight edge obliquely across the sound track slot of the analyzer. Such a wave form is shown at the top of Fig. 4. It is known that this wave form can be resolved into a series of har-

---

[7] Since a 1 per cent error in wave-length amounts to an error of about one-third of a wave in the total length of a screen of the 30th order, errors of this magnitude are quite objectionable in the higher order screens, although unimportant in the low orders.

[8] Unpublished work.

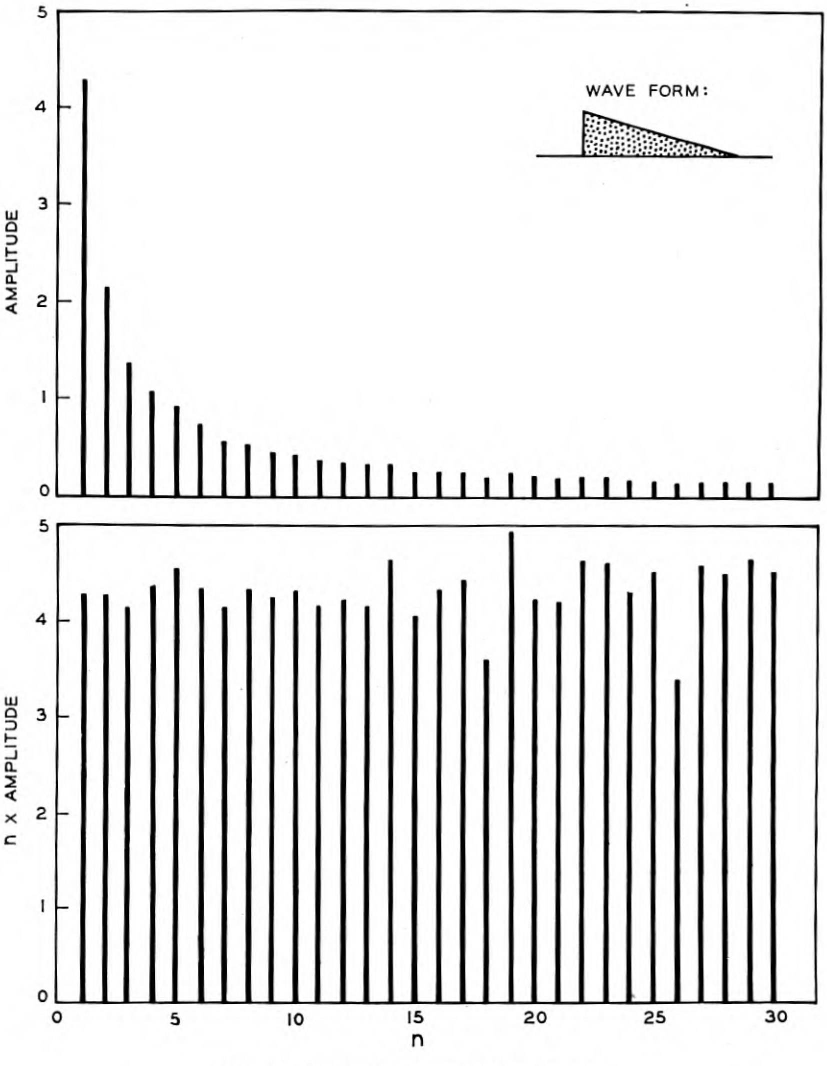


Fig. 4—Analysis of saw-toothed wave.

monics whose amplitude falls off as $1/n$ where n is the order of the harmonic. The values obtained with this analyzer are shown in the upper graph in the figure. In the lower graph each harmonic has been multiplied by n, which should make all the ordinates equal if the analysis were exactly correct.

The use of the analyzer for the sounds of speech is illustrated in Fig. 5, which shows the analyses of portions of two vowel sounds made

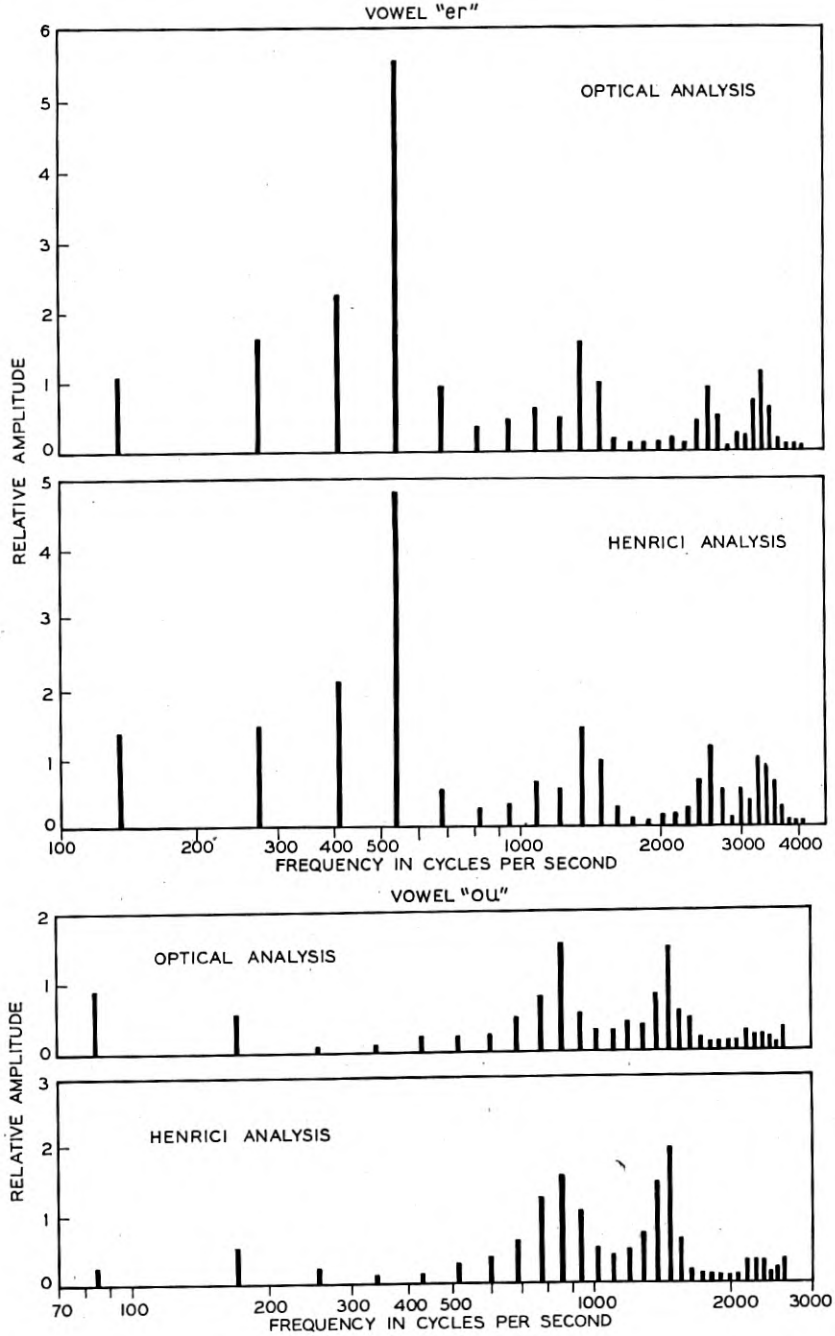


Fig. 5—Analyses of spoken vowel sounds.

with the optical analyzer. Each is compared with an analysis of the identical wave form made on a Henrici type analyzer at the State University of Iowa.[9] The first sound is a portion of the *er* in *father*. There is a very prominent fourth harmonic, indicating a strong resonance in the voice at 530 cycles. Other smaller peaks occur at 1400, 2650 and 3500 cycles. The second sound is a portion of the diphthong *ou* in *out*. It shows two peaks of about equal magnitude, with a suggestion of a third smaller one. The general features of the analyses by the two methods are seen to be in good agreement. A series of such analyses throughout the course of a spoken sound furnishes a fairly complete description of the changes in resonance, amplitude, and fundamental frequency which are taking place. Because of its high speed of operation and convenient application to records of speech on film, the present form of the optical analyzer is especially adapted to such a study of the characteristics of connected speech.

---

[9] "The Henrici Harmonic Analyzer," D. C. Miller, *Jour. Franklin Inst.*, vol. 185, pp. 285–322 (1916).

Magnetic Shielding of Transformers at Audio Frequencies

By W. G. GUSTAFSON

The first part of this article is a descriptive discussion of magnetic shielding in general. Formulae are then given for the calculation of shielding efficiency of cylindrical shells for steady and alternating magnetic fields. By means of these formulae the shielding efficiency for various types of cylindrical shields has been calculated for a steady magnetic field.

The second part of the article contains experimental information on various types of transformer shields. This information supplements the theory in connection with factors which would be very laborious to treat theoretically.

The theory and the experimental data are coordinated in such a manner that the shielding efficiency of a particular shield can be calculated with an accuracy which is sufficient for practical purposes.

IN connection with the development of repeaters for long distance telephone lines it is found that noise is introduced into the telephone lines due to magnetic pick-up by transformers and coils in the repeaters. This applies also to sound pictures equipment, public and private address systems, etc., where high gain amplifiers are used. The stray magnetic field causing this pick-up may be produced by neighboring generators, transformers, rectifiers and other power equipment. It may also be produced by other amplifier transformers and coils or by relays located in the vicinity of the disturbed coil. The intensity of the disturbing field may frequently be of the order of 0.1 oersted at the point of pick-up. However, a field intensity of the order of 0.02 oersted often causes objectionable noise and under extreme conditions values as low as 0.001 oersted may be undesirable. As the gain of the amplifier increases and the demand for good quality becomes greater, it becomes increasingly important to control magnetic pick-up. The limiting of this pick-up is in fact today one of the important problems to be considered in the design of high-gain amplifiers.

One method by which the magnetic pick-up can be decreased is by arranging the core structure and winding distribution of the transformer in such a way that the voltages induced by an external field are at least partially neutralized. In many cases, however, this impairs other important characteristics of the transformer and is therefore undesirable.

Another method is by shielding the transformer from the disturbing magnetic field. It is the object of this paper to consider such shielding and to present some data in this connection that may be of general interest.

THEORY

When a transformer is placed in an a.c. magnetic field, there will, in general, be a voltage induced in the windings. This voltage is proportional to the intensity of the magnetic field. Therefore, if the intensity of the magnetic field in the space occupied by the transformer is reduced, the induced voltage will be correspondingly reduced. This can be accomplished by enclosing the transformer in a case made of material which shields against magnetic flux. Let H_i be the intensity of the field inside the case and H_e the intensity of the field when the case is removed. The ratio H_e/H_i will then indicate the shielding efficiency of the case. Expressed in decibels:

$$\text{Shielding efficiency} = 20 \log_{10} H_e/H_i. \tag{1}$$

The shielding efficiency of the case depends primarily upon the permeability and conductivity of the material, and the mechanical construction of the case.

A high permeability material provides a magnetic path in the walls of the case of much less reluctance than the air space inside the case. The greater part of the flux will, therefore, follow the low reluctance path, and only a small part will enter the space inside the case. The higher the permeability is, the less the flux that will enter the space inside the case.* With a steady magnetic field all the shielding is due to this cause.

An alternating magnetic flux induces eddy currents in the material of the case as shown in Fig. 1. These eddy currents are a function of the conductivity and permeability of the material. They may increase or decrease the shielding efficiency of the case. That is, the eddy currents i_{e1} (Fig. 1), which are due to the component of the magnetic field perpendicular to a wall of the case, will set up a counter mmf. which will oppose flux entering the case. In a copper case, the shielding is primarily due to such eddy currents. On the other hand, the eddy currents i_{e2} (Fig. 1), which are due to the component of the field parallel to a wall of the case will set up a counter mmf.

---

* It is assumed here, of course, that the source of the magnetic flux is at some distance so that the amount of flux leaving the source is not appreciably affected by the case.

which will oppose the flux following the low reluctance path in the walls of the case, or what is the same thing, decrease the effective permeability of this path and will, in that way, decrease the shielding efficiency. In a case made of high permeability material the latter eddy currents, i_{e2}, obviously should be reduced as much as possible.

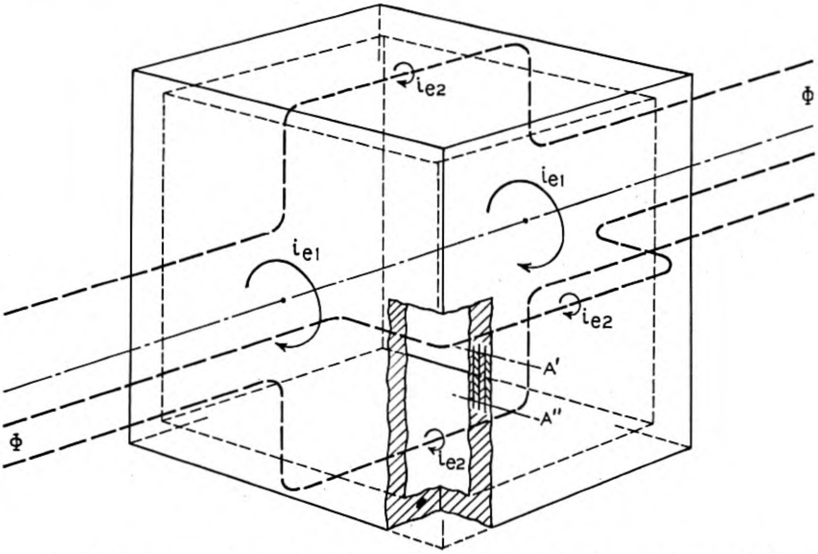


Fig. 1—Eddy currents produced in a shield by an alternating magnetic field.

If the resistivity of the material is increased, both sets of eddy currents will be decreased. If, however, the material is laminated with the sheets parallel to the wall of the case as indicated by section A'–A'' in Fig. 1, the undesired eddy currents will be reduced without affecting those which are beneficial.

The relative effectiveness of the low reluctance path and eddy currents in securing good shielding efficiency against magnetic fields depends mainly upon the frequency of the magnetic field. As a general rule, we can say that at low frequencies, the effect of the low reluctance path predominates, while the shielding effect of the eddy currents, i_{e1}, increases as the frequency increases.

This way of looking at the effect of permeability and conductivity in a magnetic shield is intended to be purely descriptive and probably would not be practicable for a mathematical treatment. It is, however, very suggestive to the design engineer.

As an illustration of the way in which the mechanical construction of the case may affect the shielding efficiency it is at once clear that the ratio between the reluctance of the magnetic path in the walls around the case and the path through the interior of the case depends upon the size of the case. Also, any openings in the case will obviously affect its shielding efficiency.

When a magnetic transformer core is placed inside the case, the reluctance through the interior of the case will decrease and the shielding efficiency of the case will also decrease. It is therefore evident that with one type of magnetic core a case might have a different shielding efficiency than with another.

To obtain general mathematical relations between the shielding efficiency and the various factors mentioned above is very difficult. By making some simplifying assumptions, however, relations can be obtained that will be useful from a practical standpoint, although they will necessarily be somewhat limited in application.

A great deal of work on the shielding efficiency of shields constructed of different magnetic materials has been done by various investigators.* Except in a few of the more recent papers,† consideration has been restricted to a steady, uniform, magnetic field where no eddy currents are produced and where, therefore, the shielding is due entirely to the magnetic properties of the material. They have also usually limited themselves to spherical and cylindrical shields. The cylinders have been considered of infinite length with the direction of the disturbing magnetic field perpendicular to the axis of the cylinder. However, with cylinders of finite length they have found that for moderate shielding efficiencies, at points inside a cylinder at a distance from the end equal to its diameter the shielding efficiency is approximately that of an infinite cylinder. The ratio between the shielding efficiency of a cylinder and that of a sphere, the radii of the two, the permeability, the thickness and construction of the walls being the same, varies from approximately 4 : 3 in favor of the sphere for very thin shells to 9 : 8 in favor of the cylinder for very thick shells. This gives some idea of how the shape of the shield affects the shielding efficiency.

Investigations by the various investigators referred to above show that the shielding efficiency of two or more concentric cylinders or spheres may be vastly greater than that of one cylinder or sphere, the amount of magnetic material being the same. They have given mathematical relations between the shielding efficiency, the permeability and the mechanical dimensions of both spheres and cylinders. Although these relations have been derived for a steady magnetic field, they may also be applied with certain limitations to an alternating

---

* See Bibliography.

† The most important exceptions are articles No. 18, 19, 21, and 28 in the Bibliography.

magnetic field. The permeability used in this case will, of course, be the effective permeability for the particular conditions and frequency under consideration. These relations refer only to that portion of the shielding effect which is independent of the eddy currents i_{e1} (Fig. 1), and the total shielding effect will, in general, be somewhat greater.

In an article in the *Physical Review* of October, 1899, A. P. Wills considers the cases of three concentric cylinders and spheres. Due to the fact that spherical shields are less suited for our purpose I will give the equations for cylindrical shields only. Wills' formula for three cylinders for large values of permeability is given by the following equation:

$$\begin{aligned} g = 1/4\,\mu\,\{(1 - q_1q_2q_3) &+ 1/16\,\mu^2 n_1n_{12}n_2n_{23}n_3 \\ &+ 1/4\,\mu[(n_1n_3 + n_1n_2 - n_1n_2n_3)n_{12} \\ &+ (n_1n_3 + n_2n_3 - n_1n_2n_3)n_{23} - n_1n_3n_{12}n_{23}]\} + 1. \quad (2) \end{aligned}$$

In this equation

$$g = \frac{H_e}{H_i}, \tag{3}$$

where H_e is the density of the magnetic field at a point P with the shield removed and H_i is the density of the magnetic field at that point when enclosed by the shield. μ is the permeability of the material at the frequency in question. We have

$$\begin{aligned} q_1 &= r_1^2/R_1^2, & n_1 &= 1 - q_1, \\ q_2 &= r_2^2/R_2^2, & n_2 &= 1 - q_2, \\ q_3 &= r_3^2/R_3^2, & n_3 &= 1 - q_3, \\ q_{12} &= R_1^2/r_2^2, & n_{12} &= 1 - q_{12}, \\ q_{23} &= R_2^2/r_3^2, & n_{23} &= 1 - q_{23}, \end{aligned} \tag{4}$$

where r_1, R_1, r_2, etc., are the various radii of the cylinders as shown by Fig. 2.

By making $q_3 = 1$ (or $n_3 = 0$), in (2), we get the relation for two concentric cylinders.

$$g = 1/4\,\mu(1 - q_1q_2 + 1/4\mu n_1n_2n_{12}) + 1. \tag{5}$$

By making $q_2 = 1$ (or $n_2 = 0$), (5) changes into an equation for one shell only

$$g = 1/4\,\mu(1 - q) + 1. \tag{6}$$

It has been shown (A. P. Wills, *Phys. Rev.*, vol. 24, page 243, February 1907) that for a shield of predetermined size, that is when the smallest and the largest radii (r_1 and R_3 in the case of three cylinders, Fig. 2) are specified, the radii of the surfaces of the successive cylinders should be in a geometric progression to give the most efficient shield. That is, we should have $q_1 = q_2 = q_3 = q_{12} = q_{23} = q$. Equation (2) then becomes

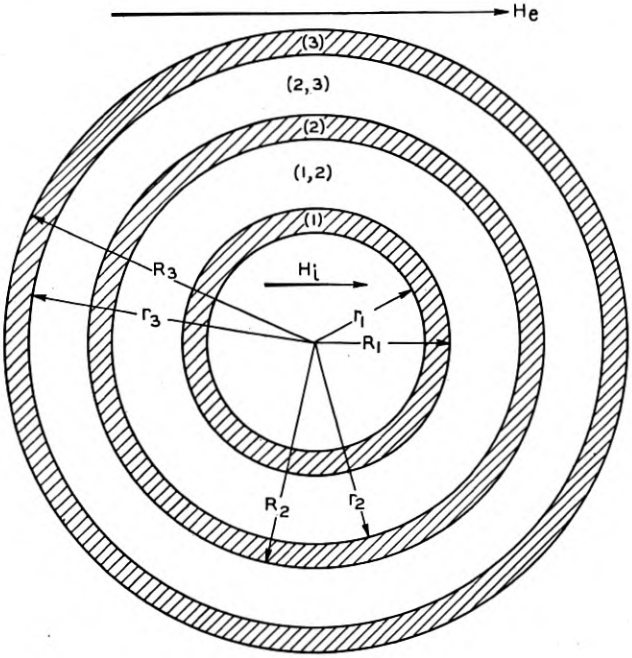


Fig. 2.

$$g = 1/4\,\mu(1 - q^3 + 1/16\,\mu^2 n^5 + \mu n^3 + 3/4\,\mu n^4) + 1. \tag{7}$$

For two cylinders we get

$$g = 1/4\,\mu(1 - q^2 + 1/4\,\mu n^3) + 1. \tag{8}$$

In these equations, the following relations hold between r_1 and R_3 for (7) and r_1 and R_2 for (8)

$$R_3 = r_1/\sqrt{q^5}, \qquad R_2 = r_1/\sqrt{q^3}. \tag{9}$$

The effect upon the shielding efficiency of varying s (Fig. 3) from zero to P keeping $r_1/R_1 = r_2/R_2$ can be obtained by means of the following equation

$$g = 1/4\,\mu\left[1 - \frac{k^2}{q_{12}} + 1/4\,\mu\left(1 - \frac{k}{\sqrt{q_{12}}}\right)^2 n_{12}\right] + 1, \tag{10}$$

where

$$k = \frac{r_1}{R_2}.$$

If R_1/r_2 (that is $\sqrt{q_{12}}$) is varied from 1 to R_2/r_1 the desired result is obtained.

Assume in Fig. 3 the thickness of the two cylinders to be the same, that is, $R_1 - r_1 = R_2 - r_2$. The variation in shielding efficiency vs. s

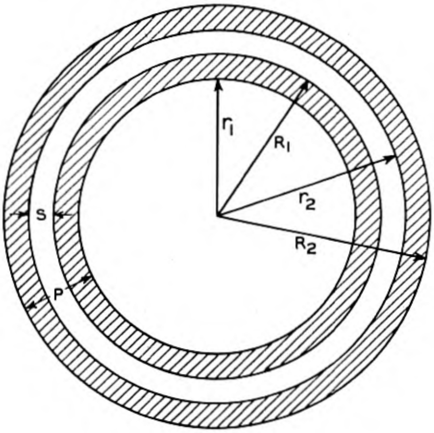


Fig. 3.

or $\sqrt{q_{12}}$, is then given by equation (5), with q_2 expressed in terms of q_{12} and q_1 as follows:

$$\sqrt{q_2} = \frac{1}{1 + \sqrt{q_{12}}[1 - \sqrt{q_1}]}. \tag{11}$$

In an article in the *Philosophical Magazine* of February 1933 L. V. King has developed relations for the shielding efficiency of spherical and cylindrical shells taking into account the effect of induced currents. The following equations for an infinitely long metallic cylinder have been picked from his paper. For a non-magnetic shell, the thickness of which is small compared to its radius, the shielding ratio, g, is given by

$$g = |\cosh(ka) + 1/2\, ka \sinh(kd)|, \tag{12}$$

where a is the radius and d the thickness in cm. k is given by

$$k = 2\pi\sqrt{\frac{f}{\rho \cdot 10^9}}(1 + i), \tag{13}$$

where f is frequency in cycles and ρ is resistivity in ohms per centimeter cube. At low frequencies (12) reduces to

$$g = \left| 1 + i\frac{4\pi^2 fad}{\rho} 10^{-9} \right|, \tag{14}$$

which is good up to about 10^5 cycles. The direction of the disturbing magnetic field in (12) and (14) has been assumed perpendicular to the axis of the cylinder.

Other formulæ which take into account both conductivity and permeability are also given in King's article. They are, however, rather complicated and require elaborate calculations.

Mr. S. A. Schelkunoff in an article in the October 1934 issue of the *Bell System Technical Journal* has derived formulæ which are comparatively simple although they take into account both conductivity and permeability. His treatment is quite different from that presented above and his results are expressed in terms of radial impedances. For an infinitely long cylindrical shield the diameter of which is large compared to the radial thickness of the shield the shielding efficiency, S, is given by

$$S = R + A. \tag{15}$$

In this formula R is the sum of the reflection losses at the surfaces of the shield. We have

$$R = \sum_{n=1}^{n} 20 \log_{10} \frac{|k_n + 1|^2}{4|k_n|} \text{ db}, \tag{16}$$

where k_n is the ratio of the radial impedance in the first medium to that in the second. That is,

$$k_n = \frac{Z_1}{Z_2}. \tag{17}$$

The radial impedance for a good dielectric is given by

$$Z = 2\pi f \mu \rho i \text{ ohms}. \tag{18}$$

For a metal

$$Z = \sqrt{\frac{2\pi f \mu i}{g}}. \tag{19}$$

In (18) and (19) f is the frequency in cycles, ρ is the radius in cms., g is the intrinsic conductance in mhos/cm., and μ is the intrinsic inductance in henries/cm.*

A in (16) is the sum of the attenuation losses in the successive shells. For any one shell

$$A = 8.686\alpha t \text{ db}, \tag{20}$$

where $\alpha = \sqrt{\pi g \mu f}$ and t is the radial thickness in cms.

Calculated Curves

By means of the equations (2) to (11) the shielding efficiency of various types of cylindrical shields has been calculated. The permeability considered is 5000 which is readily obtainable at low frequencies and field strengths by means of permalloy.† The calculations are given in the form of curves in Figs. 4, 5 and 6.

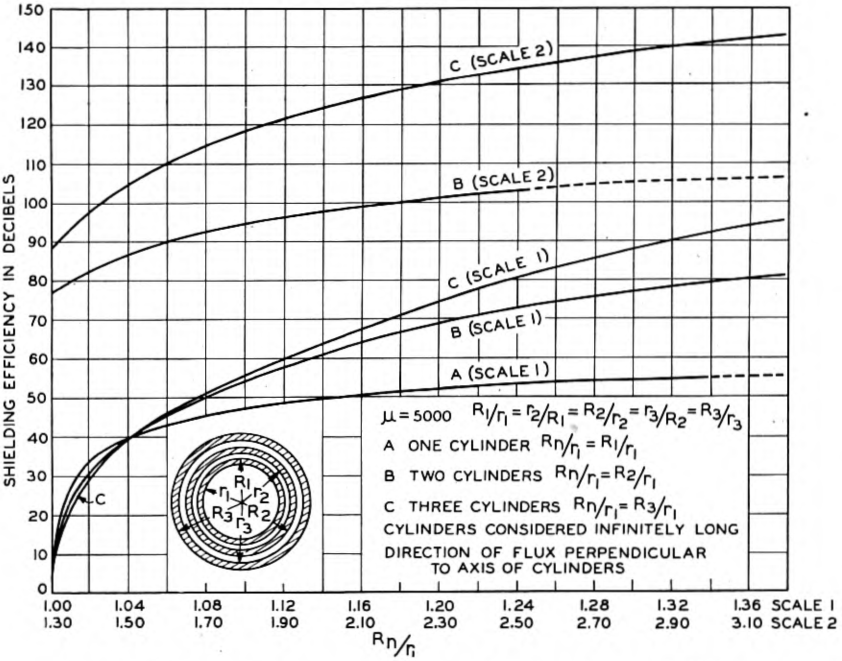


Fig. 4—Shielding efficiency of one, two, and three concentric cylinders (for zero frequency).

---

* Thus in empty space (or dielectrics, approximately) $\mu = 4\pi 10^{-9}$ henries/cm. In general $\mu = 4\pi\mu_0 10^{-9}$ where μ_0 is the intrinsic permeability referred to empty space as unity.

† Arnold and Elmen, "Permalloy," *Journal Franklin Institute*, May, 1923, pp. 621–632.

The curves of Fig. 4 show the shielding efficiency of one cylinder and of two and three concentric cylinders with air space between. The shielding efficiency is given as a function of the ratio between the outside and inside radii of the shield, that is, R_n/r_1, where R_n is the outside radius of the outside cylinder and r_1 is the inside radius of the inside cylinder. These curves show the relative shielding efficiencies of 1, 2 and 3 cylinders and give numerical values for μ equal to 5000. The relative dimensions of the cylinders are such that the ratios between the inside and the outside radii of the cylinders and of the air spaces between the cylinders are in geometric progression. These curves show that when a very high shielding efficiency is desired it is not only advantageous but necessary to use two or more cylinders. Thus with one cylinder the maximum shielding efficiency that can be considered practical when $\mu = 5000$, is approximately 50 db. The maximum theoretical limit for one, two and three cylinders is 62 db, 124 db and 186 db respectively when $\mu = 5000$.

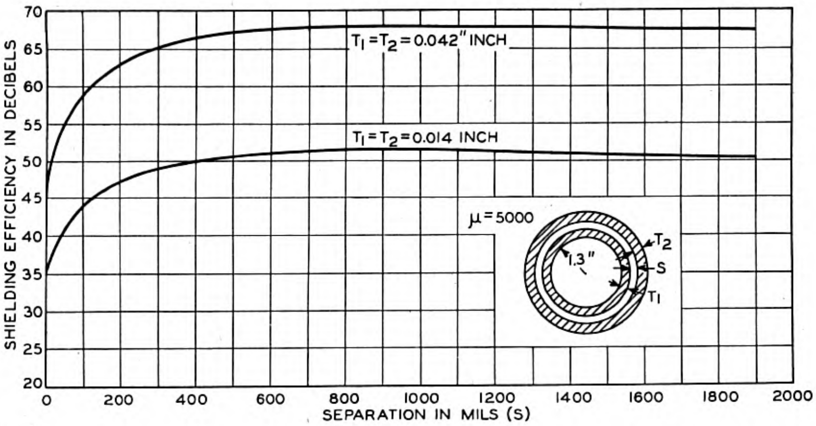


Fig. 5.—Shielding efficiency of two concentric cylinders versus the air-gap between them. Thickness of the wall of each cylinder kept constant. Zero frequency.

In Fig. 5 is shown the shielding efficiency of two concentric cylinders vs. the thickness of the air space between the cylinders. The thicknesses of the walls of the two cylinders are equal. Two thicknesses of the walls of the cylinders have been considered, namely, .014″ and .042″. An interesting fact is brought out by comparing the curves of Fig. 5 with curve "*B*" of Fig. 4. For example, when the air space is .042″, the upper curve of Fig. 5 gives a shielding efficiency of two cylinders with an air space between them, the thickness of which is the same as that of the cylinders. The ratio between the outside radius and the inside radius is 1.097. The shielding efficiency of this combination is approximately the same as that of two cylinders having the same ratio R_n/r_1 as given in Fig. 4, curve "*B*." This shows that the condition that the radii of the cylinders should be in geometric progression is not very critical, at least, not for the value of R_n/r_1 under consideration.

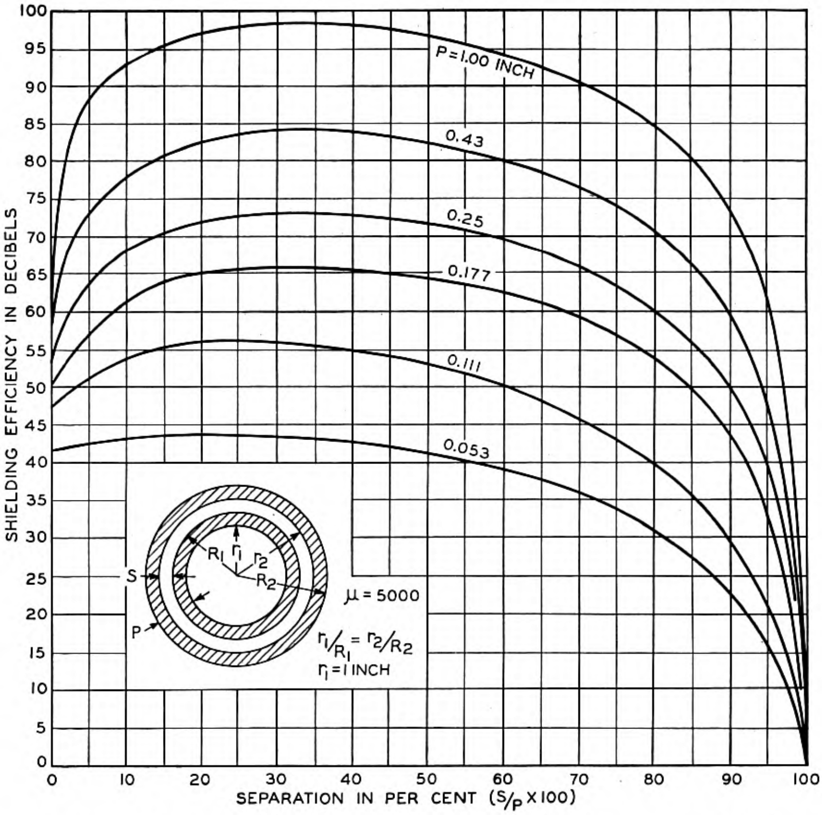


Fig. 6—Shielding efficiency of two concentric cylinders versus the air-gap between them. Overall thickness of the wall of the shield (*P*) kept constant. Zero frequency.

The curves of Fig. 6 also show the shielding efficiency of two concentric cylinders vs. the air-gap between them. In this case, however, the total thickness *P* of the wall of the double cylindrical shield is kept constant and the air-gap is increased from zero to *P* at the expense of the cylinders. When the air-gap is zero, the shielding efficiency is therefore that of a solid cylinder, the thickness of the wall of which is P. When the air-gap is equal to P, the thickness of the two cylinders is zero, and hence the shielding efficiency is zero. The air-gap is so located that the radii of the two cylinders are in geometric progression.

EXPERIMENTAL DATA

From the discussion under "Theory" it is evident that although the equations (2 to 11 incl.) were derived with the assumption of a steady magnetic field the above calculated values apply equally well to an alternating magnetic field if the effective permeability is used. However, the results are far from sufficient to determine the shielding efficiency of a magnetic shield for a transformer. The shielding due to the eddy currents, i_{e1} (see Fig. 1), is not included. To include this the formulæ (12) to (20) inclusive must be used. The equations have been derived with the assumption that the length of the cylinders is infinite. In practical applications this is obviously not so. Covers, however, approximately counterbalance the effect of the finite length of the cylinders. The magnetic core of the transformer also materially affects the shielding efficiency. In connection with such factors as these which would be very laborious to treat theoretically some experimental information will now be given. The frequency range of the disturbing magnetic field was limited to from 50 to 4000 cycles.

The shielding efficiency has previously been defined as follows:

$$\text{Shielding Efficiency} = 20 \log_{10} H_e/H_i. \tag{1}$$

From the standpoint of the shielding of transformers we are primarily interested in the reduction of the transformer terminal voltage which is caused by the disturbing magnetic field. For this reason it will be found convenient to define the shielding efficiency in connection with transformers in decibels as follows:

$$\text{Shielding Efficiency} = 20 \log_{10} E_e/E_i, \tag{21}$$

where E_e is the terminal voltage due to the disturbing magnetic field with the shield removed and E_i the corresponding voltage with the transformer inside the shield. In addition, E_e and E_i are restricted to the maximum terminal voltages, with respect to position, that is, the transformer is assumed to be in that angular position with respect to the direction of the magnetic field, in which the maximum terminal voltage is obtained. With an unshielded shell type transformer, for example, this position would be that in which the axis of the winding coincides with the direction of the disturbing magnetic field. This restriction is necessary for the definition (21) to be of any value.

With a small air core coil and an infinite cylinder the axis of which is perpendicular to the axis of the winding and to the direction of the disturbing magnetic field the two definitions are equivalent.

The circuit used in making measurements consists of a field coil producing a magnetic field, a pick-up coil and a vacuum tube voltmeter. The field coil is a loop two feet in diameter consisting of 500 turns of wire. The magnetic field at the center of this coil is uniform over a space sufficiently large for the purpose. The pick-up or search coil when placed in the magnetic field will have a voltage produced across its terminals. This voltage is measured with the vacuum tube voltmeter and E_e and E_i in equation (21) are thus obtained. The size of the shell type core on which the pick-up coil is wound is 3″ × 1 29/32″ × 1″ where the 1 29/32″ dimension is parallel to the axis of the winding, the 1″ dimensions being the pile-up of laminations. Unless otherwise mentioned this coil was used in all of the following measurements.

Permalloy having an initial permeability of the order of 5000 at low frequencies was employed for both the core and the shield throughout this investigation.

Permalloy Cases

In Fig. 7 is shown the shielding efficiency vs. frequency of a rectangular permalloy case which consists of five contiguous layers of .014″ thick permalloy sheet. The size of this case is approximately 2 3/8″ × 2 1/2″ × 3 1/4″ and the position of the coil inside the case is such that the axis of the winding is parallel to the 2 1/2″ dimension. The relative size of the coil and case is such that there is approximately

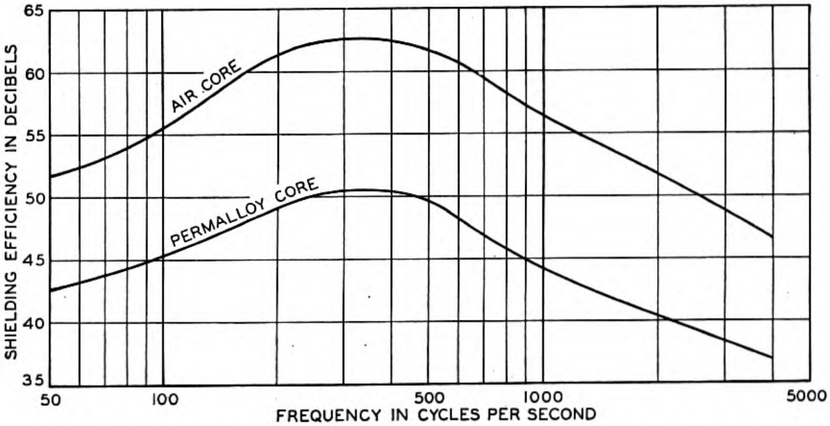


Fig. 7—Observed shielding efficiency of laminated permalloy case.

1/8″ clearance between the core of the coil and the case. The shielding efficiency is given for a coil having a permalloy core and also for a coil having a non-magnetic core. These curves illustrate the effect of the magnetic core upon the shielding efficiency. At low frequencies the shielding is mainly due to the magnetic properties of the shield material. As the frequency increases, however, the effect due to the eddy currents, i_{e1} (Fig. 1), increases and the shielding efficiency increases. At approximately 300 cycles a maximum is reached and from here on up to 4000 cycles the shielding efficiency decreases. This is due to the fact that at these frequencies the eddy currents, i_{e2} (see Fig. 1), decrease the effective permeability of the material at a greater rate than the shielding efficiency is increased due to the eddy currents i_{e1}. The slope of the curves at 50 cycles is not zero. This shows that even at 50 cycles there is a considerable shielding effect due to the beneficial eddy currents.

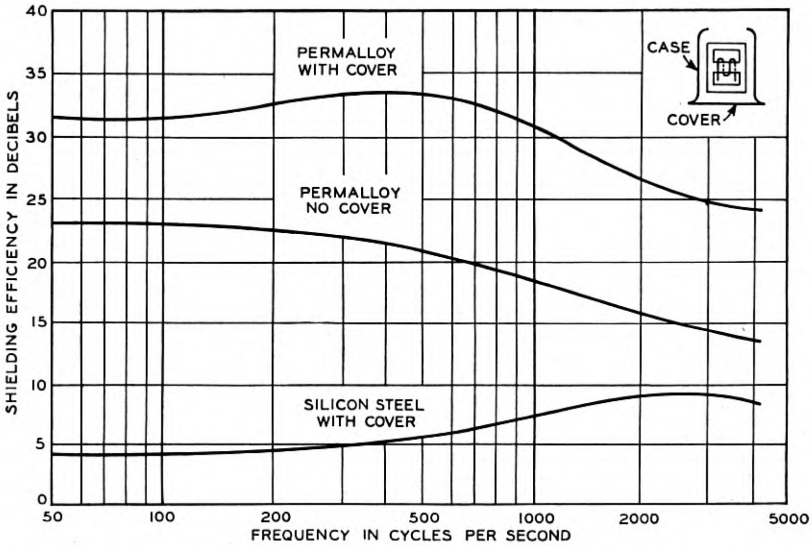


Fig. 8—Observed shielding efficiency of cases made of 1/32″ thick permalloy and silicon steel sheet.

The curves of Fig. 8 show the shielding efficiency vs. frequency of a cylindrical permalloy case, the thickness of the walls of which is 1/32″. The approximate dimensions of this case are 3 1/4″ high × 2 5/8″ diameter and the relative dimensions of the case and core are such that there is approximately 1/8″ clearance between the core and the case. A comparison between the two curves for permalloy shows the effect of the cover upon the shielding efficiency.

A comparison between the curve for permalloy and that for silicon steel on Fig. 8 gives a striking example of the advantage of using permalloy instead of steel.

Effect of Air-Gap Between Core and Shield

It has been pointed out previously that the magnetic core decreases the reluctance through the interior of the case and in that way decreases the shielding efficiency. When the size of the core is such as to almost fill the case, that is, when the air-gap between the core and the case is small, this effect becomes large. In Fig. 9 are given some data in this connection for cylindrical shields. Six permalloy cylinders were used in this illustration. They are so constructed as to fit over each other, the smallest fitting snugly over the transformer core ("d" = 0). Each cylinder consists of two layers of .014″ thick permalloy sheet. The cylinders have been numbered 1 to 6 from the smallest to the largest, respectively. Curve "A" of Fig. 9 shows the shielding efficiency with a non-magnetic core and curve "B" gives the corresponding information with the permalloy core having a permeability of approximately 5000 at 70 cycles and low field strengths which are the conditions under which the measurements were made.

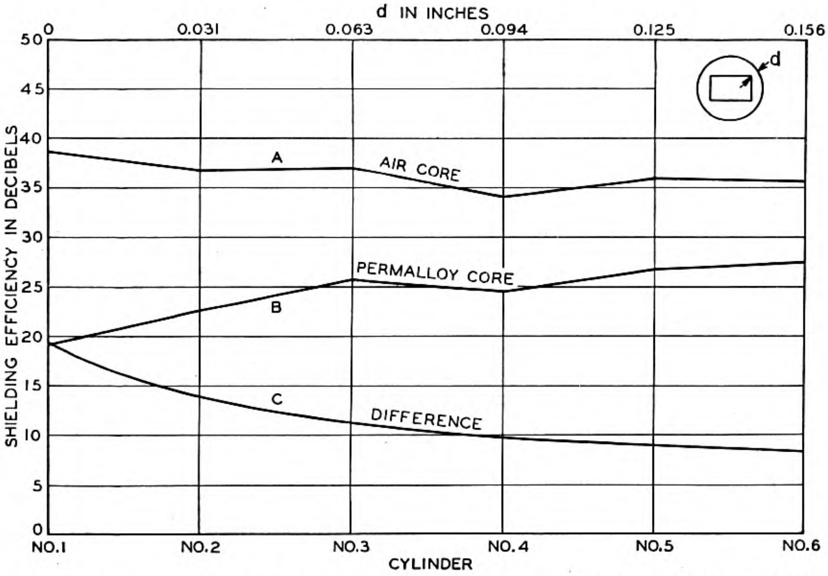


Fig. 9—Observed effect of air-gap between core and shield. Cylindrical shield. Frequency 70 cycles.

These curves have been drawn discontinuously because the shielding efficiency is a function not only of "*d*" but of other factors such as permeability, size of the cylinders, etc. Curve "*C*" on the other hand is primarily a function of "*d*" and has, therefore, been drawn continuously. This curve gives the difference between the shielding efficiency with an air core and with a permalloy core and shows the advantage of increasing the air-gap between the core and the cylinder. Thus, for example, in this particular case approximately 8 db better shielding efficiency is obtained with an air-gap of 1/16″ than if there is no air-gap.

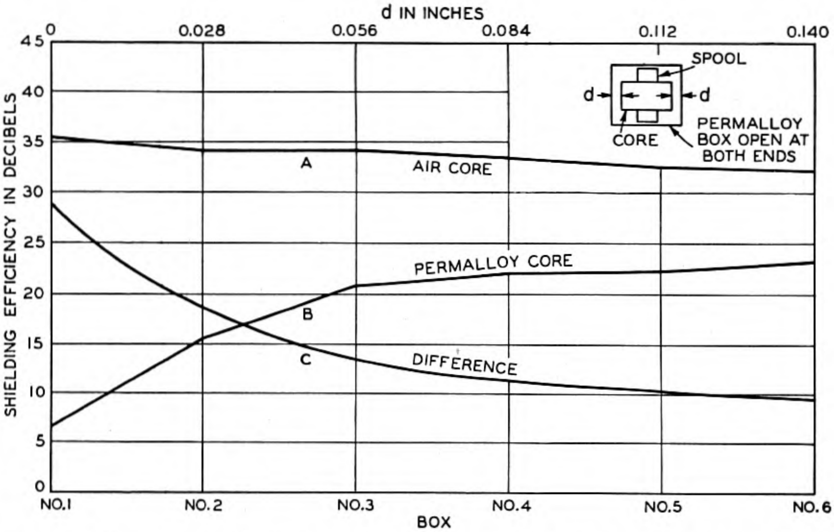


Fig. 10—Observed effect of air-gap between core and shield. Rectangular shield. Frequency 70 cycles.

Similar curves are given in Fig. 10 for rectangular boxes of the same height and same wall thickness as the cylinders. The measurements were made at 70 cycles per second. Due to the larger contact area between the core and the shield the effect of the air-gap is much greater here than with cylinders.

As the effective permeability of the walls of the shield increases the effect of an air-gap increases, other things being equal. In general it may also be said that as the shielding efficiency increases (due to increased thickness of the walls, for example) the effect of an air-gap increases.

If there is an appreciable air-gap between the core and the shield a large variation in the effective permeability of the core will affect the shielding efficiency very little. That is, practically the same results will be obtained with a silicon steel core, having a permeability of 400 as with a permalloy core having a permeability of 5000. On the other hand if the silicon steel core is replaced by an air core the change in the shielding efficiency will be of such an order as is indicated by Fig. 7. The reason for the small effect of changing from a permalloy core to a silicon steel core as compared to the changing from a silicon steel core to an air core is, of course, due to the fact that in the first case there is a decrease in permeability of 12.5 : 1 while in the second case the corresponding reduction in permeability is 400 : 1.

High Efficiency Shields

A shielding efficiency of from 20 to 50 db is, for many purposes, sufficient in connection with the shielding of transformers. Occasions arise, however, when a shielding efficiency much greater is desired. To accomplish this by means of a simple case, magnetic material having a permeability much greater than is now readily available would be needed. The curve *B* of Fig. 11 gives the shielding efficiency

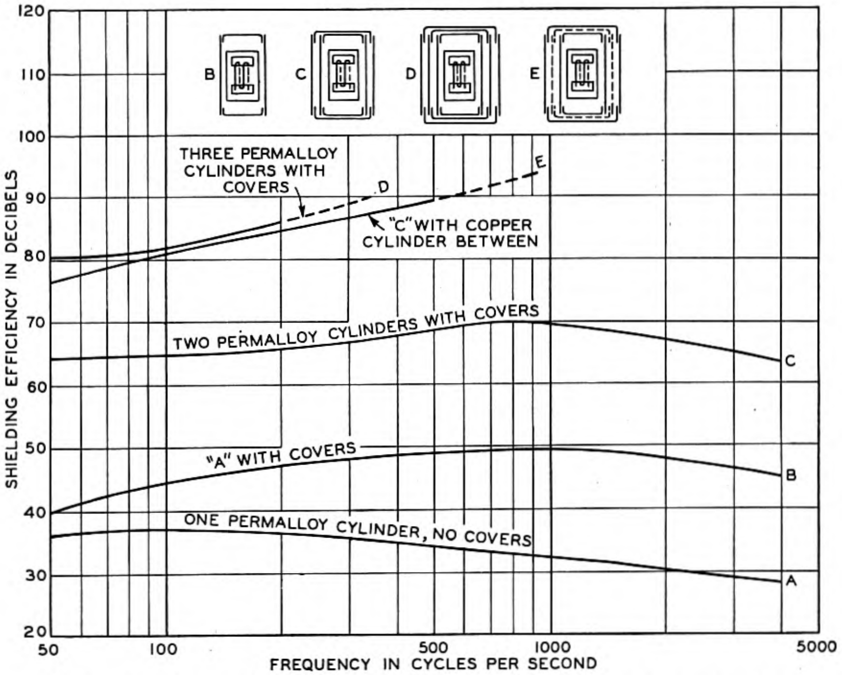


Fig. 11—Observed shielding efficiency of various high-efficiency shields.

of a permalloy cylinder which has a permeability of approximately 5000. This cylinder is 4.5″ high, inside diameter 2.5″, and thickness of wall equal to .07″. By increasing the thickness of the wall the efficiency would be only slightly increased. This is evident from a study of Fig. 4. However, by placing a second cylinder over "*B*" a substantial improvement is obtained (Curve *C*). Still greater shielding efficiency is obtained by placing a third cylinder over the former two as shown by curve *D*. The dimensions of the second and third cylinders are such that the ratios between the outside and inside radii of the three cylinders and of the air-gaps between them are approximately in geometric progression. The height of the second and third cylinders is also 4.5″ and the effective permeability at low frequencies and field strengths approximately 5000.

Since the effective permeability of the permalloy used in the above shields is close to 5000 we can compare these data with the theoretical curves of Fig. 4, which were calculated with the permeability assumed equal to 5000. This comparison shows that the theoretical analysis of the shielding of infinite cylinders against steady magnetic fields may be applied to the shielding of transformers. Due to such factors as a magnetic core inside the shield, eddy current shielding, end effects etc. only an approximate check can be expected. At 50 cycles per second the measured values for one cylinder and for two and three concentric cylinders are 40, 64, and 80 db respectively. Corresponding calculated values as given by Fig. 4 are 41.5, 66, and 89 db respectively.

It is evident from the effect of the copper cylinder between two permalloy cylinders as shown by curve *E* (Fig. 11) that three permalloy cylinders with copper cylinders between the inner and middle and between the middle and outer will give a shielding efficiency of the order of 100 db (voltage ratio $E_e/E_i = 10^5$) from 50 to 4000 cycles.

Effect of Covers

The information given in Fig. 12 shows the importance of covers. This figure gives the shielding efficiency of a cylinder which consists of two layers of .014″ permalloy sheet. Curve *A* gives the shielding efficiency without any covers and curve *B* shows the advantage of adding covers which overlap 1/2″ and consist of two layers of .014″ permalloy sheet. The two curves *C* give the shielding efficiency of the same cylinder provided with flat plate covers, C_1 representing covers .014″ thick and C_2 covers .028″ thick. The relative size of the coil and cylinder is such that there is a clearance of approximately 1/16″ between the core and the magnetic shield. It is evident that in this particular instance the covers are very important. Regarding

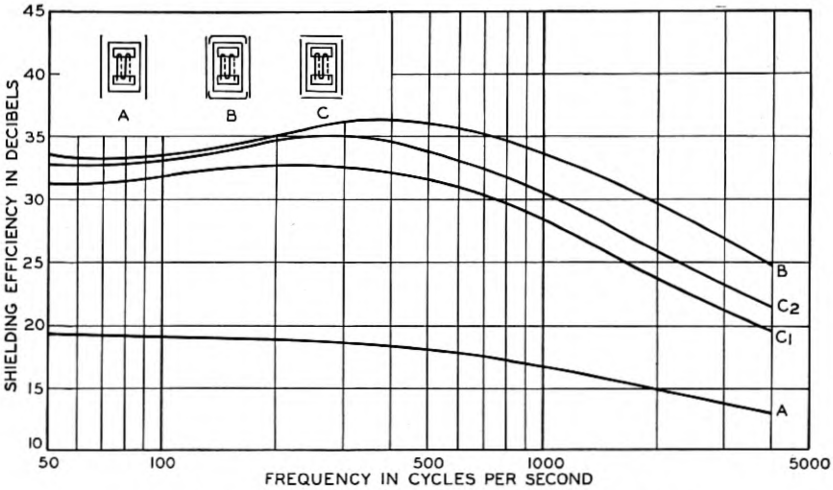


Fig. 12—Observed effect of covers on cylindrical shields.

contact between the cover and the cylinder it may be shown that at low frequencies this is immaterial while at higher frequencies the reverse is true. However, even at these higher frequencies a small overlap (as for *B*) is sufficient.

Examples of the effect of covers upon the shielding efficiency are also given by Figs. 8 and 11. The effect of covers on a copper cylinder as shown by curves *A* and *B* (Fig. 13) is of special interest.

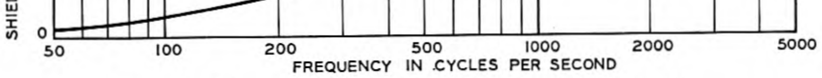


Fig. 13—Observed shielding efficiency of a copper cylinder.

Use of Copper

The curves *A* and *B* of Fig. 13 give the shielding efficiency vs. frequency of a cylinder made of copper. This cylinder has an inside diameter of 2 5/8″ and is 4.5″ high. The thickness of the wall is 1/16″. The shielding efficiency is here entirely due to the eddy currents, i_{e1} (Fig. 1). Curve *B* shows that, after the effect due to the open ends of the cylinder is eliminated by means of covers, the shielding efficiency is approximately proportional to the logarithm of the frequency.

Although copper has a very low shielding efficiency at low frequencies when used alone, tests show that under certain conditions it is very effective when used in conjunction with permalloy. This is illustrated by a comparison between the curves *C* and *E* of Fig. 11. The copper cylinder is similar to the one referred to above except that the thickness of the wall is only 1/32″. The permalloy cylinders are those for which the shielding efficiency is given by curve *C* of the same figure.

Another striking example of the use of copper in conjunction with permalloy is furnished by Fig. 14. The curve *A* gives the observed shielding efficiency of a permalloy box when there is an airspace between the core and the box of 1/32″ ($d = 1/32''$). If 1/32″ thick copper plates are inserted between the core and the box (see Fig. 14), there is a great improvement in the shielding efficiency as a comparison between the curves *A* and *B* shows. This improvement is 10 db at 50 cycles although the effect of the copper plates alone would be of the order of one db, as is evident from the curves on Fig. 13. At higher frequencies the improvement is still better. It is approximately 15 db between 100 and 4000 cycles. Approximately the same results

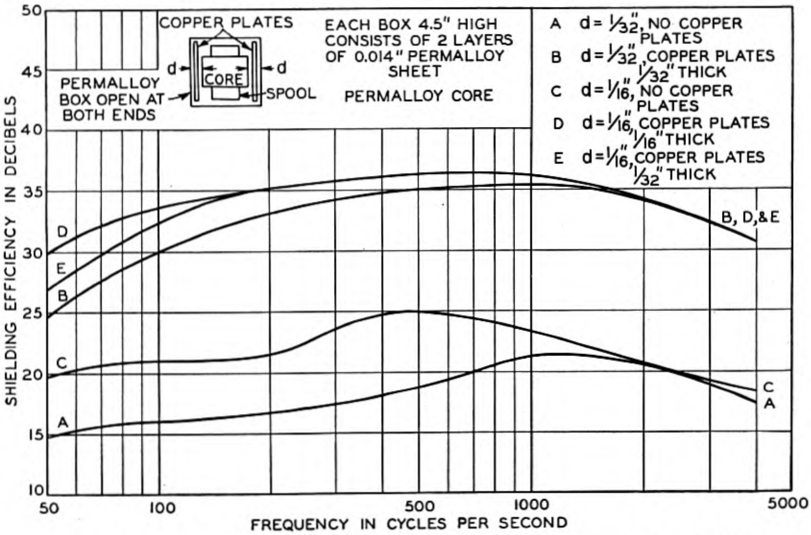


Fig. 14—Observed effect of copper between core and permalloy shield.

are obtained with a spacing of 1/16″. A comparison between the curves C and E shows the improvement which is obtained in this case with 1/32″ copper plates. The effect is somewhat less than with a spacing of 1/32″, at least at frequencies below 1000 cycles. At low frequencies copper plates, 1/16″ thick, show a slight improvement over the 1/32″ copper plates as is shown by the curve D.

Curve B on Fig. 10 shows that if the airspace between the core and the box is small the shielding efficiency is very low. In a case like this a copper spacer is very effective. For example, a 5 or 10-mil copper plate replacing an airspace of the same thickness will greatly improve the shielding efficiency.

GENERAL

The foregoing is a discussion of the magnetic shielding of transformers from external magnetic fields. The reverse problem of shielding a transformer or coil so as to prevent its magnetic field from affecting other apparatus has not been considered. However, it is safe to assume that approximately the same degree of shielding will be obtained, provided the leakage field does not produce excessive saturation in the shield. That is, assuming that a power transformer is producing a disturbing magnetic field in the space occupied by an input transformer, then a shield over the power transformer will produce approximately the same effect as a shield over the input transformer, where each shield has been constructed in accordance with the information on the foregoing pages. This has been demonstrated experimentally in an article by J. E. R. Constable in the *Wireless World* of February 26, 1937.

Although this paper has been restricted to the magnetic shielding of transformers it is equally applicable to any apparatus which is susceptible to inductive pick-up. This is because in any apparatus where there is inductive pick-up there is in effect a coil. It may be an actual coil and it may be only a loop of lead wires.

The author wishes to thank Mr. E. T. Hoch for many helpful suggestions.

BIBLIOGRAPHY

1. J. Stefan, *Wied. Annalen*, Vol. 17, p. 928, 1882.
2. A. W. Rücker, *Phil. Mag.*, Vol. 37, p. 95, 1894.
3. H. DuBois, *Wied. Ann.*, Vol. 63, p. 348, 1897; Vol. 65, p. 1, 1898, also *Electrician*, Vol. 40, 1897, pp. 218, 316, 511, etc. cont.
4. A. P. Wills, *Phys. Rev.*, Vol. 9, p. 193, 1899; Vol. 24, p. 243, 1907.
5. James Russell, *Roy. Soc., Edinb., Trans.*, Vol. 40, p. 631, 1903.
6. E. F. Nichols and S. R. Williams, *Phys. Rev.*, Vol. 27, p. 250, 1908.
7. W. Esmarch, *Ann. der Phys.*, Vol. 39, p. 1540, 1912.
8. W. W. Coblens, *Bul. Bu. of Standards*, Vol. 13, p. 423, 1916.

9. C. Benedicks, *Ann. der Phys.*, Vol. 72, p. 236, 1923.
10. R. H. Barfield, *Inst. Elec. Eng.*, Vol. 62, p. 249, 1924.
11. D. W. Dye, *Jour. Sci. Inst.*, Vol. 3, p. 65, 1925.
12. A. V. Hill, *Jour. Sci. Inst.*, Vol. 3, p. 335, 1925.
13. J. H. Morecroft and A. Turner, *I. R. E., Proc.*, Vol. 13, p. 477, 1925.
14. H. Pleijel, Svenska Ingeniörsvetenskapsakademiens Handlingar, NR 49, 1926.
15. H. L. Curtis, *A. I. E. E. Journal*, Vol. 48, p. 453, 1929.
16. S. L. Gokhale, *A. I. E. E. Trans.*, Vol. 48, p. 1307, 1929.
17. N. Hillers, *Telefunken Zeitung*, Vol. 13, pp. 13–28, 1932.
18. L. V. King, *Phil. Mag.*, Vol. 15, p. 201, February 1933.
19. W. Lyons, *I. R. E., Proc.*, Vol. 21, p. 574, April 1933.
20. G. W. O. Howe, *Wireless Engr. and Exp. Wireless*, Vol. 11, p. 347, July 1934.
21. S. A. Schelkunoff, *Bell Sys. Tech. Jour.*, Vol. 13, p. 532, October 1934.
22. *Radio Engg.*, Vol. 15, p. 11, July 1935.
23. T. E. Sterne, *Rev. Sci. Inst.*, Vol. 6, p. 324, October 1935.
24. W. F. Randall, *The Nickel Bulletin*, Vol. 9, No. 4, p. 73, April 1936.
25. R. Bachstroem, *Arch. f. Elektrotechnik*, Vol. 30, p. 267, April 1936.
26. S. A. Schelkunoff, *Radio World*, Vol. 29, p. 47, April 1936.
27. Samuel Levy, *I. R. E., Proc.*, Vol. 24, p. 923, June 1936.
28. C. W. Oatley, *Phil. Mag.*, Vol. 22, p. 445, September 1936.
29. J. E. R. Constable, *Wireless World*, Vol. 40, p. 198, February 1937.
30. W. Herzog, *E. N. T.*, Vol. 14, p. 81, March 1937.

Coaxial Cable System for Television Transmission*

By M. E. STRIEBY

THE reports which have been made on the progress in television development increase the expectation that the broadcasting of visual programs will soon be realized. In anticipation of that result, the Bell Laboratories has been engaged for some time in the development of wire line circuits for transmitting television signals between studios and broadcasting transmitters, or between cities, as may some day be required if television follows in the footsteps of sound program broadcasting.

The wide frequency bands required for television and the dearth of available frequencies appear to force the broadcasting of television signals into the ultra-high frequency range. At these high frequencies, the coverage which can be obtained from a broadcasting station is very limited as compared to that obtainable in the sound broadcasting frequency range. Hence, if television programs are to reach large sections of the country simultaneously, the provision of interconnections between large numbers of television broadcasting transmitters will become even more important than it is today for sound broadcasting.

Coaxial cables have received much publicity as transmission lines for television. The original conception and use of the coaxial form of cable was first as a low frequency submarine conductor and later as a lead-in for radio antennas. The idea of a coaxial cable or other medium for the transmission of very broad frequency bands originated in the course of telephone development in America.[1] The first lengths of such cable for broad-band transmission were made here and its first use for the transmission of a large number of simultaneous telephone conversations was between New York and Philadelphia.[2] In this country the important reason for developing coaxial cable systems was, and still is, that they appear to offer material economies in the provision of large groups of long distance telephone facilities. Television has been secondary.

Recently experiments have been made on the transmission of television signals over the coaxial cable between New York and Philadelphia. This cable contains two coaxial conductor units within

---

* Presented at A. I. E. E. Winter Convention, New York City, Jan. 27, 1938. Published in June 1938 issue of *Elec. Engg.*

a lead sheath about 7/8″ in diameter as indicated in Fig. 1. Two coaxial units were provided because, for long distance telephone operation four-wire operation is preferable, one coaxial being employed for transmission east to west and the other west to east. Each coaxial unit is made up of a 13-gauge inner conductor on which hard rubber disks have been placed at intervals of 3/4 of an inch. The outer metallic tube is made up of 9 overlapping copper tapes so designed that they form essentially a solid copper tube about 20 mils in thickness and .267″ in inside diameter.

The transmission loss of this circuit as a function of frequency is shown on Fig. 2 together with the portion of the attenuation that is contributed by conductance losses. Inasmuch as the intention is to use these conductors at very high frequencies, a high grade insulating material was used with the result that the conductance losses are small. It should be noted that the attenuation increases very nearly as the square root of the frequency.

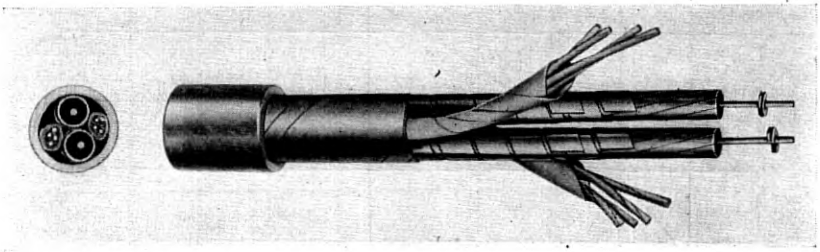


Fig. 1—Section of the New York-Philadelphia coaxial cable.

In order to transmit high frequencies over long distances, a great deal of amplification is obviously required. The New York-Philadelphia cable was initially equipped to handle a band of one million cycles. Its overall attenuation at a million cycles is approximately 600 decibels. In order to reduce this to a usable amount 10 repeater points were provided at intervals of about 10 miles each having an amplification at the top frequency of about 60 decibels. These repeaters were so designed that they provided less gain at low frequencies than at the high frequencies, in approximately the same degree as the line had less attenuation. To make up for certain cumulative irregularities an equalizer was built and added to the overall circuit. The net result was a transmission path which had approximately zero loss over the whole band which it was desired to use from 60 kc. to 1000 kc. as shown in Fig. 3.

Another complicating factor is, however, involved, as is the case with all long wire circuits, namely, the variation encountered with changes in temperature. The loss in a 10-mile repeater section varies materially from summer to winter. If the cable is hung overhead, this variation is about as shown in Fig. 4 and amounts to a change of about ± 7 per cent in the attenuation. If the line is buried underground at the normal depth used for telephone cables, the actual variation is about one-third as much. Automatic transmission regulators were developed to compensate for these changes. These regulators depend, for their operation, on the transmission of a pilot channel. At each point where it is desired to regulate the transmission, the pilot channel is selected by a very narrow band filter and its amplitude used to control an automatic device which changes the gain of the repeater until the amplitude of the pilot at the output of the repeater reaches a certain predetermined value. The regulators on the New York-Philadelphia circuit have operated with such accuracy that it has been unnecessary to make manual adjustments to take care of temperature changes.

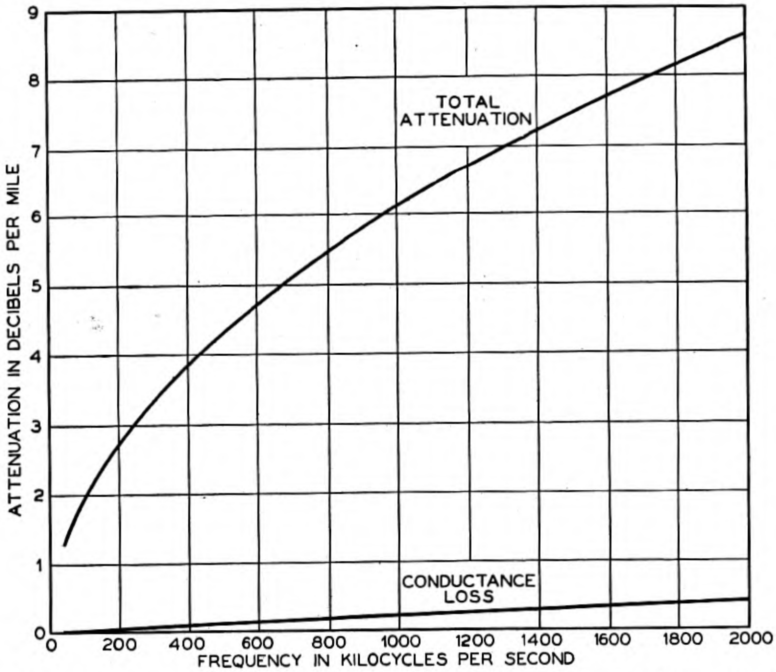


Fig. 2—Attenuation per mile of the New York-Philadelphia coaxial cable. Also the proportion of that attenuation due to conductance losses.

The repeaters used along this route were novel in that most of them were placed in small iron boxes located at convenient points along the line. Power for their operation was transmitted at sixty cycles over the coaxial cable itself. Figure 5 shows one such unattended repeater located near Dunellen, New Jersey.

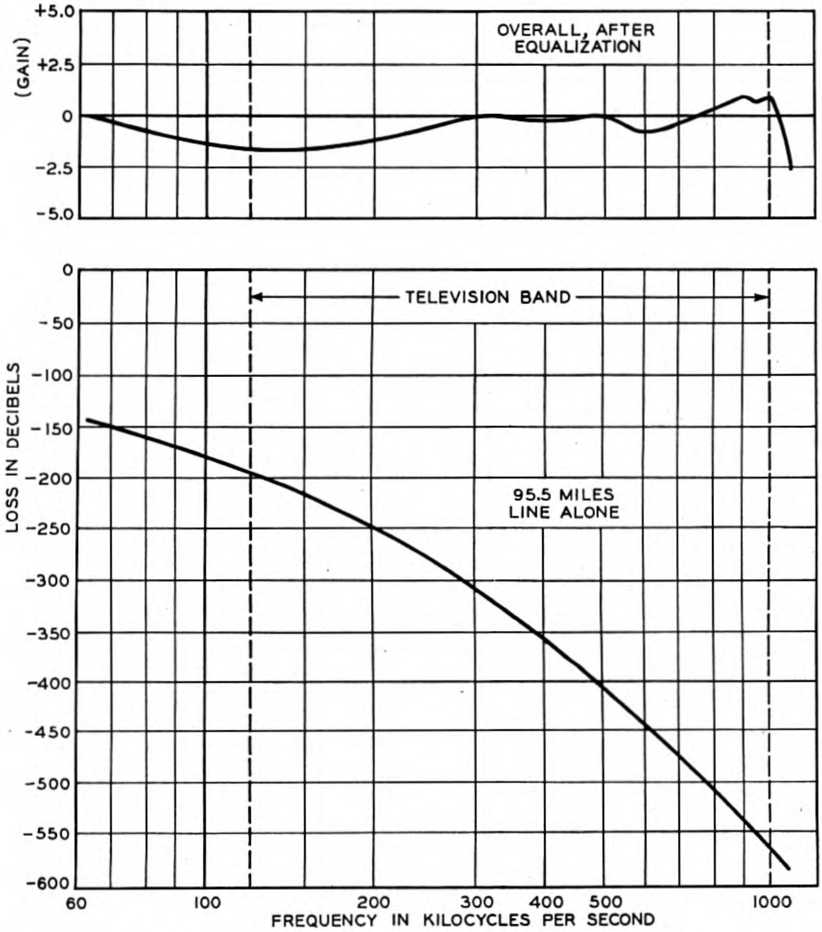


Fig. 3—Overall attenuation of the New York-Philadelphia cable without repeaters and the net attenuation after repeaters and equalizers had been provided.

The coaxial system between New York and Philadelphia was designed to provide 240 telephone circuits. Skeleton terminal apparatus was installed at New York and Philadelphia to test out such operation. This apparatus has been described [2, 3, 4] in various papers and will not be discussed here. Suffice it to say that methods and equipment have been developed which enable a wide band to be split up so as to obtain hundreds of telephone channels.

For television transmission a quite different problem existed—namely, can very wide band systems be used for the long distance transmission of these complicated signals? In planning tests to be significant of the operation of the cable system for such signals, it was important to obtain as nearly as possible an ideal television signal and as nearly as possible an ideal television receiver. In this way it was hoped that any defects in the cable transmission itself could be discovered.

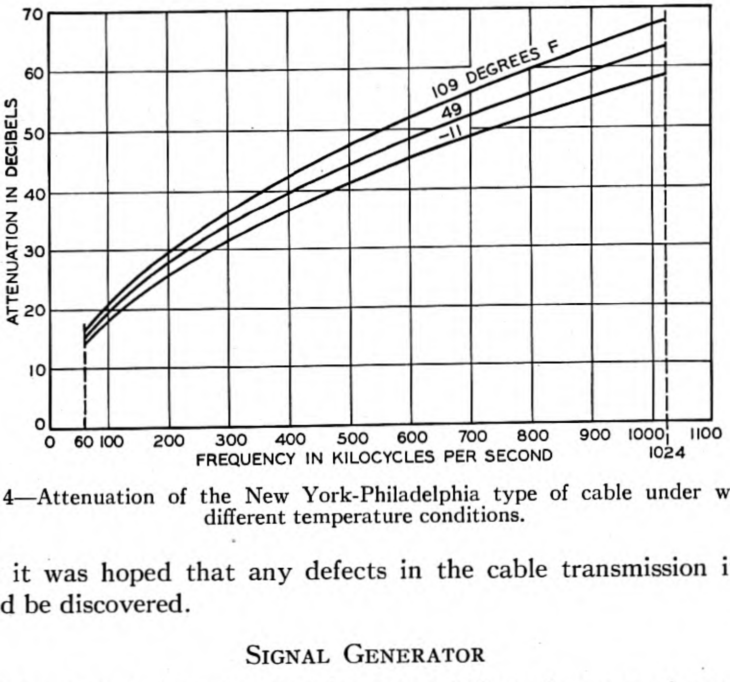


Fig. 4—Attenuation of the New York-Philadelphia type of cable under widely different temperature conditions.

Signal Generator

Although television implies the transmission of an actual scene it is much more satisfactory for engineering studies to transmit a motion picture, since exactly the same picture can then be transmitted over and over again as the circuit elements are changed or adjusted. Moreover, it was decided to use mechanical scanning to obtain the most nearly perfect signal possible, and with this form of scanning a film can be much more brightly illuminated than an actual scene and hence is much easier to use. Because of these factors a motion picture film was chosen as the material for the recent experiments.

The scanning disk used in these tests was developed under the direction of Dr. H. E. Ives at the Bell Telephone Laboratories. It consists of a six-foot disk with a circle of 240 lenses near its outer edge.

The arrangement is indicated schematically in Fig. 6. Light from a powerful incandescent lamp behind the disk, passing through one lens at a time, is focussed by the lens to form on the film a small dot of light about three thousandths of an inch square. The lenses in the disk are spaced by a distance equal to the width of the picture, or a little less than an inch, and as the disk rotates, each spot is moved rapidly across the picture. The film is carried at a uniform rate


Fig. 5—Repeater near Dunellen, New Jersey.

downward behind the disk at such a speed that the successive holes throw their light in successive rows across the picture one above another. The film moves one frame for each revolution of the disk. A photosensitive surface mounted behind the film picks up the light transmitted through it, and produces a complex electric current corresponding to the variations of light in the picture. Figure 7 is a photograph of the housing in which the disk is mounted. This scanning arrangement produced a picture of 240 lines, 24 frames per second. It was recognized that 24 frames per second were not sufficient to avoid flicker but this choice simplified the scanning apparatus and it was believed would not interfere with engineering tests.

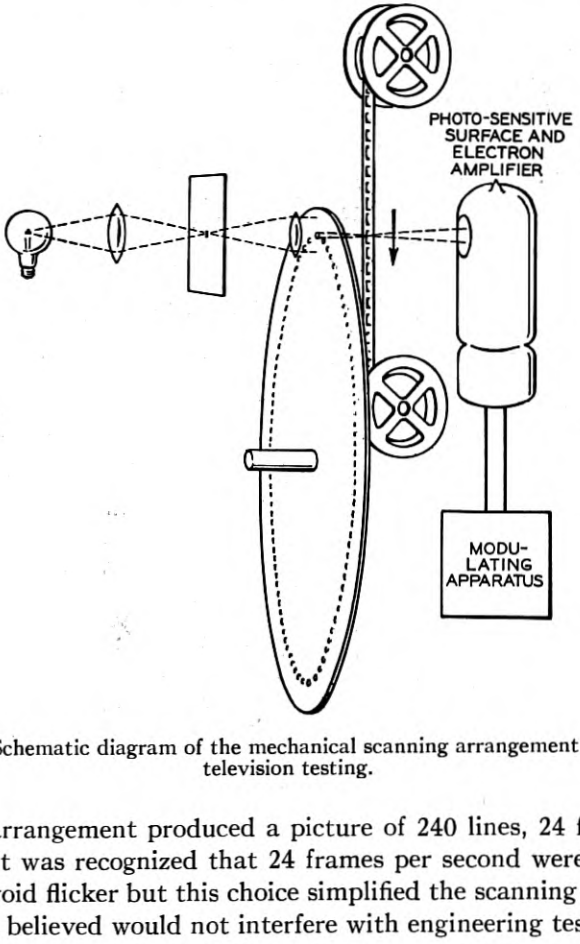


Fig. 6—Schematic diagram of the mechanical scanning arrangement used for television testing.

Signal Frequency Range

In order to understand what frequency is required to transmit an image scanned in this way consider the diagram shown in Fig. 8.

Of course, actual television will not ordinarily deal with such a picture, but by means of it an approximate visualization of the problem can be obtained. A desired definition of 240 lines was chosen and it was decided that the requirement would be to transmit square picture elements as indicated in the figure. The shape of the picture which it was convenient to use with this scanning disk is wider than it is high in the ratio of 7 : 6. This differs somewhat from the standard aspect ratio of 4 : 3, but is easily taken into account. The total number of picture elements in a frame is then $240^2 \times \frac{7}{6} = 67200$.

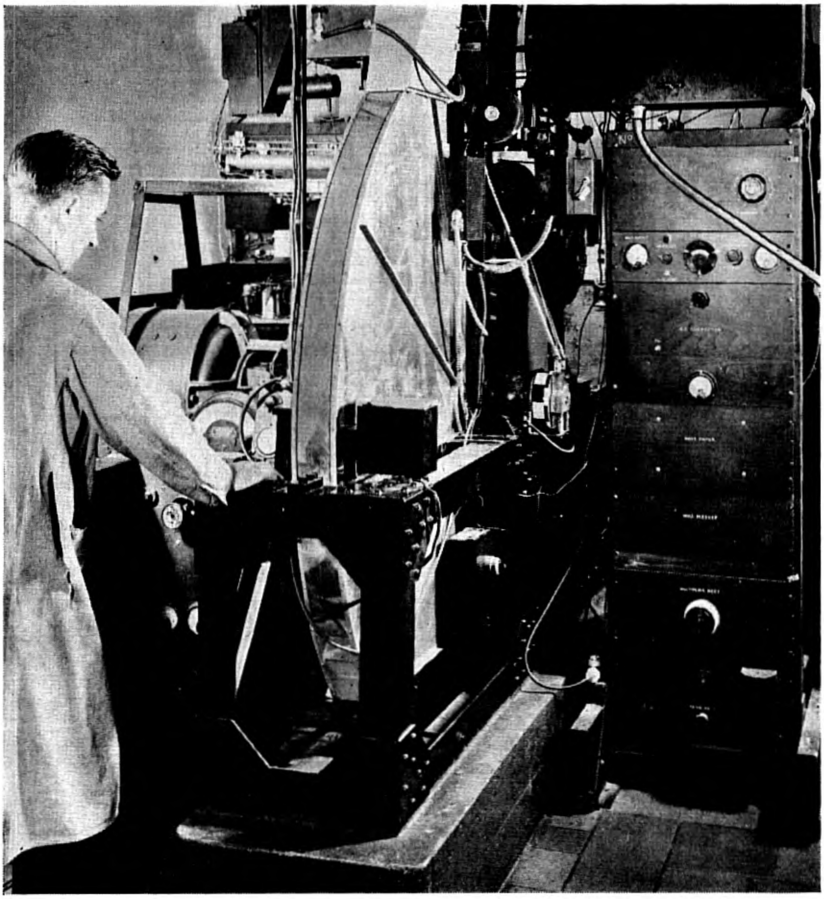


Fig. 7—Photograph of the mechanical scanning disk and associated experimental apparatus.

If the smallest picture element to be transmitted is a single block then the distribution of light and shade over the block is unimportant. The average brightness over the block is what counts. Obviously, a simple approximation is a sine wave as shown at the bottom of the picture. This wave has ½ cycle for each block and is as high a frequency as there is any profit in transmitting for this diagram.

The top frequency needed for such a picture can then be calculated. The number of square elements in the picture computed above is 67,200. As each of these elements represents ½ cycle, this figure is divided by 2, giving 33,600 cycles per frame. As similar frames are reproduced 24 times a second, the result is 24 × 33,600 or 806,400 cycles per second. In a real moving picture, other frequency components may exist at all other frequencies from 800 kc. down to and including direct current. The direct current or zero frequency component controls the general level of brightness of the picture. Where the general level of brightness changes slowly, it results in a component of very low frequency. A composite picture can be

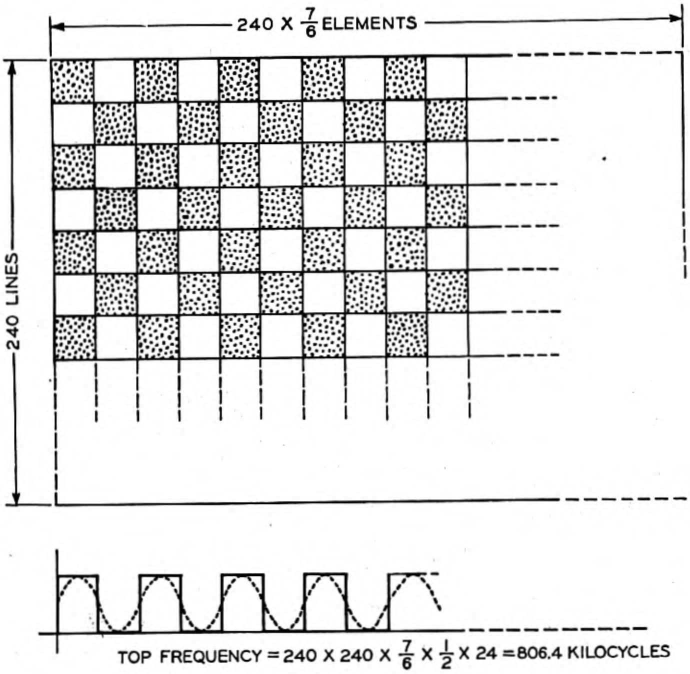


Fig. 8—Diagram illustrating the resolution of an image into picture elements and the derivation of maximum frequency required for transmission.

imagined which will produce a pronounced component at any given frequency, hence, it is deemed important to transmit the entire band from 0 to 806 kc.

RECEIVING DEVICE

At the receiving end an effort was made to obtain as high a degree of fidelity of reproduction as possible. No small factor in the success of the recent experiment was the special cathode ray tube, designed by Dr. C. J. Davisson and used at the receiving end to display the transmitted picture. Some of the features of this tube are indicated schematically in Fig. 9. A stream of electrons from the cathode passes through a series of electron lenses which focus a narrow beam on an aperture .006″ square. Between the lenses and the aperture, however, are two modulating plates connected to the incoming circuit in such a way that there appear on these plates potentials varying according to the voltage of the incoming signals. The effect of potential on these plates is to deflect the electron beam, and the conditions are such that at maximum strength of signal practically the entire stream of electrons passes through the hole and forms a brilliant spot of light on the front of the tube. As the signal decreases in strength, the electron stream is more and more deflected; less electrons pass through the aperture, and the illumination on the sensitized end of the tube decreases.

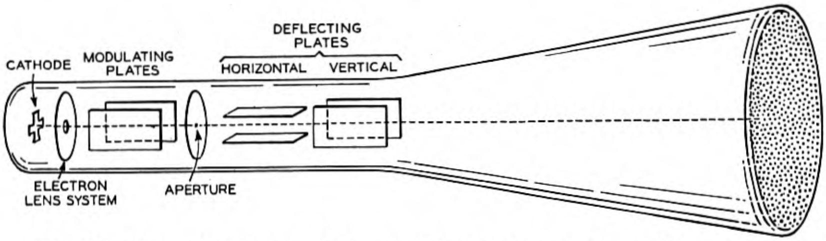


Fig. 9—Schematic diagram of the special cathode ray tube used for viewing transmitted images.

In addition to these modulating plates, and placed between the aperture and the front of the tube, are two other pairs of plates mounted in planes at right angles to each other. The potential on one of these sets of plates, controlled by a frequency of 5760 cycles, which is the frequency at which successive lines are scanned, varies in such a way that the beam of electrons passing through the aperture is swept across the front of the tube from left to right, exactly in synchronism with the scanning beam at the sending end. After the beam reaches the farther side of the picture, the potential on the plates is suddenly changed, and the beam is rapidly moved back to begin the next line. Due to a black mask down the far side of the film being scanned, there is no signal during this very short period while the voltage on the plates is changed, and thus the electron beam is deflected from the aperture and is not visible on the front of the tube during its return.

The potential on the other pair of plates is controlled at a frequency of 24 cycles per second, which is the rate of scanning successive frames. The effect of the potential on these plates is to deflect the electron beam downward in synchronism with the motion of the film at the sending end. This results in the passage of the electron beam across the front of the tube in successive rows, one below another. After the last row has been scanned, the voltage on the plates is changed and returns to the value that causes the beam to appear at the top line of the tube. A properly synchronized blanking-out pulse is introduced between successive frames of the film, so that no signal is received during this interval, and thus the passage of the electron beam from the bottom to the top of the frame is not visible.

Figure 10 is a photograph of one of these cathode ray tubes. Due to its superior design, the image is very sharp over the entire field and a wide range of brightness is secured. The chief factors in its success are the sharp focusing by the electron lenses, the linear deflection of the beam at the aperture, and the great length of the tube, which makes it necessary to deflect the electron beam over only a narrow angle to cover the 7 × 8 inch field. Since this trial was a test to determine the capabilities of the system, such matters as size and cost, which would be important with commercial receivers, were not controlling.

Modulation System

The frequency band which was generated at the sending end as noted above was 0 to 806 kc. The coaxial cable system used could not transmit this band, because repeaters were not designed to pass frequencies below about 60 kc. This limitation was incorporated in the original design because the cable offers insufficient shielding to various disturbances at low frequencies. For television transmission it was necessary, therefore, to raise the television signal band to a higher frequency position before attempting transmission over the line. A number of considerations led to the decision to raise the entire frequency band 144 kc. for transmission over the coaxial cable.

Where such a wide frequency band is to be raised by an amount less than the width of the band itself, a single modulation is not generally satisfactory. The products of modulation include the original frequency band as well as the upper and lower sidebands, so that there will be a confusing jumble of frequencies in the modulator output. For this reason a double modulation method was used for the recent experiments.

Fig. 10—Photograph of one of the special cathode ray receiving tubes.

The modulating scheme employed can be followed with the help of Fig. 11, which shows the two modulating steps at the sending end and the two demodulating steps at the receiving end in four lines beginning at the top. A carrier of 2376 kc. is used for the first modulation, which results in a lower sideband from 1570 to 2376 kc. and an upper sideband from 2376 to 3182. The carrier itself is eliminated in the balanced modulator. The output of this modulation is passed through a filter, but because the two sidebands touch each other at 2376 kc., the filter cannot cut off all the upper sideband. At the output of this filter there is thus the lower sideband plus a small amount of the lower part of the upper sideband. The upper sidebands from all subsequent modulations are readily eliminated by filters because of the wide separation.

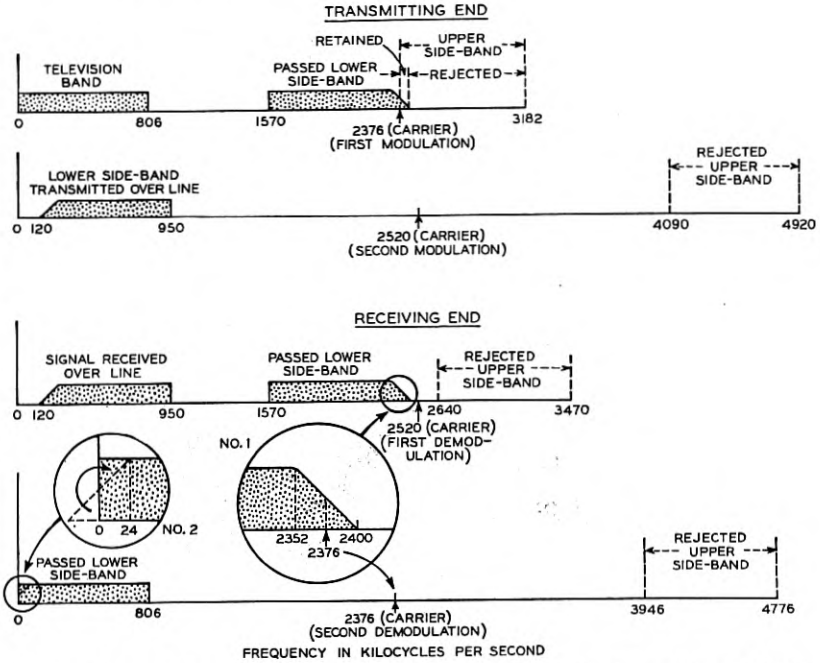


Fig. 11—Schematic diagram illustrating the processes of modulation and demodulation used in transmitting television signals over the coaxial line.

The carrier for the second modulation is 2520 kc., and the lower sideband extends from 950 down to 144 kc. plus a vestigial upper sideband remaining from the first modulation which extends down to 120 kc. The high-pass filter following this modulation is accurately designed to pass with controlled attenuation a group of frequencies just above 144 kc. and the vestigial sideband. The resulting single sideband, extending from 120 to 950 kc., is then passed over the coaxial cable.

At Philadelphia the received signal, together with a carrier of 2520 kc., is applied to the first demodulator, and the lower sideband, from 2400 down to 1570, is passed to the second demodulator where a carrier of 2376 kc. is applied. The lowest frequency of the lower sideband, 1570 kc., is converted to 806 kc., becoming the highest frequency of the final demodulated band. The frequencies from 2352 to 2400 kc. of the sideband before the second demodulation have been attenuated somewhat by the high-pass filter following the second modulator, and the second demodulating carrier, 2376 kc., falls in the middle of this attenuated band as shown in inset No. 1. Frequencies extending about 24 kc. above the carrier are inverted by the demodulation, and superimposed upon the corresponding frequencies just below the carrier. The magnitude and phase of these components are proportioned by the high-pass filter and an equalizer so that the overall result, when they are superimposed, is an essentially flat transmission band from 0 to 806 kc.

The above steps of modulation involved a number of difficulties. In the first place the signal level must be carefully controlled so that on the one hand it does not sink into the background noise, while on the other hand it must not be raised to such high levels that unwanted modulation products are produced in too great magnitude. The first modulator presents some special problems. It must accommodate all frequencies from 0 to 806 kc. In order to eliminate the carrier, it must be balanced to a very high degree—about 80 db in this case. The reason the carrier must be so completely wiped out is that the 0 frequency component of the signal is identical with the carrier at the output and hence the true d-c. value of the signal must be exceptionally free from carrier interference.

Referring to Fig. 12, the next piece of apparatus is a band filter to eliminate the video signal and cut off the top edge of the band. Then follows the 2nd modulator which is quite conventional. A low-pass filter is next and is very important as it performs part of the function of cutting off and adjusting the vestigial sideband. Then follows an amplifier, a predistorting network to partially equalize the amplitudes of the different components of the signal, an aperture equalizer to correct for the fact that the scanning spot is of finite size, a terminal equalizer to make up for irregularities in the overall setup and other amplifiers.

The carrier apparatus at the sending end is shown on Fig. 13. It is mounted in rather conventional form except for the 1st modulator which was arranged to minimize the effect of low-frequency vibrations. At the receiving end about the same apparatus is required in the inverse order and will not be discussed in detail.

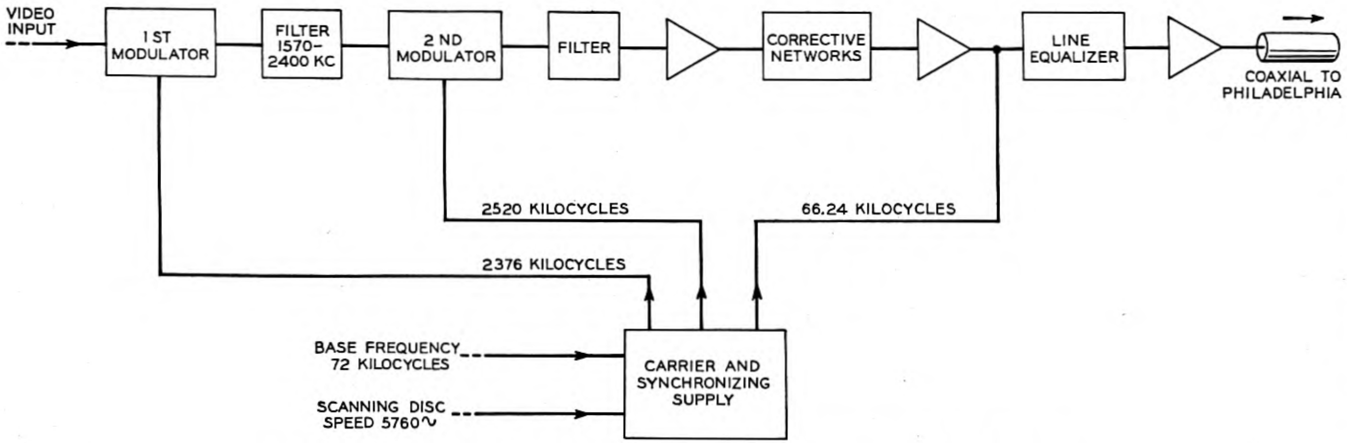


Fig. 12—Block diagram of the transmitting carrier television terminal at New York.

CARRIER SUPPLY

Another important feature of this system was the provision of accurately spaced carriers for the various modulating operations. The Bell System 4 kc. standard was used to produce a 72 kc. (18th harmonic) base frequency signal which could be transmitted over the line and so tie the transmitting and receiving ends together. From this, by harmonic generation, the carriers used in the two steps of modulation were produced, being the 33rd and 35th harmonics.

Fig. 13—Photograph of the transmitting carrier television terminal with associated sound apparatus.

To synchronize the scanning at the receiving end with the transmitting disk required another direct tie. The disk is driven by a d-c.

motor and its speed cannot be kept very constant. By means of an auxiliary lamp and photocell, the frequency with which one scanning line followed another—approximately 5760 per second—was obtained. This was modulated with the 72 kc. mentioned above and transmitted over the line as a lower sideband at 66.24 kc. At the receiving end by demodulation, the exact line speed is obtained and used to drive the horizontal sweep. The vertical sweep is obtained by generating the 240th subharmonic of this—namely 24 c.p.s.

LINE EQUALIZATION AND TEST RESULTS

Returning now to the line transmission problem, the signals which might be transmitted in the general case are indicated in Fig. 14. They include the pilot channels used for automatic transmission regulation at 60 kc. and 1024 kc. For convenience, one telephone channel with a carrier at 64 kc. is indicated as an order wire, and a wide-band program channel with carrier at 84 kc. to transmit the sound. These, of course, could be provided with ordinary telephone facilities. The base frequency of 72 kc. and the disk synchronizing sideband at 66.24 are also included. For the television signal itself the band from 120 to 950 kc. is provided. Actually in the tests to Philadelphia, automatic regulation was not needed and a separate wire line was used for synchronization.

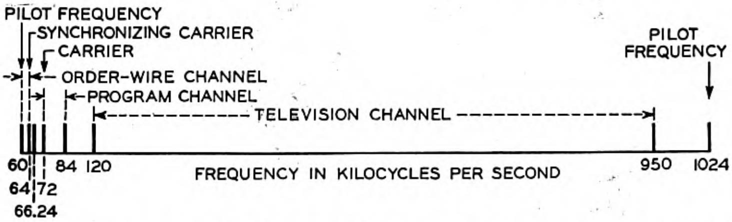


Fig. 14—Frequency allocation for a television transmission system with associated control circuits.

It was necessary to provide networks and equalizers to insure that the coaxial line did not distort the ultimate image due to unequal attenuation, resulting in amplitude distortion, or to unequal time of transmission, causing phase distortion. The actual attenuation characteristics of the line[5] and the overall result were shown above in Fig. 3. The requirements for phase distortion are rather difficult to meet. The details in the scanned picture result in various frequencies of the electrical signal, and if these details are to appear in the reproduced picture in the same relative position as in the scanned picture, it is essential that all frequencies be received in very closely the same relative time relationship as they are generated. Referring back to Fig. 8, it was assumed that no picture element could be displaced by more than about half its width. This led to the decision to hold frequencies between 806,000 and 3000 cycles to a delay distortion of about 0.3 microsecond. For a similar degradation of detail in the vertical direction, the permissible delay distortion is 280 times as great which, in a system of this type, is very easily obtained. The actual circuit roughly met these requirements as indicated by Fig. 15, which shows the phase delay characteristics of the line, repeaters and equalizers, and of the overall circuit at the frequencies used for transmission.

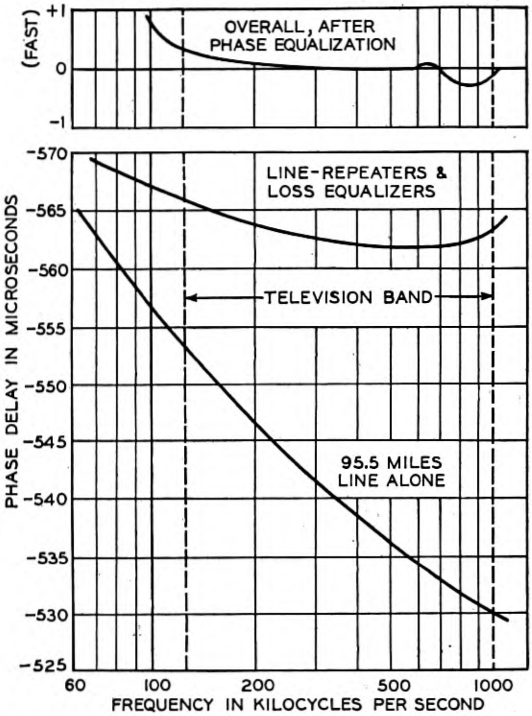


Fig. 15—Phase delay of New York-Philadelphia television circuit.

Noise or interference is very annoying in television transmission; and pattern, or single frequency interference, is particularly objectionable. The permissible noise or interference depends on the amplitude range of the reproduced picture. During these experiments, it was found that a substantially linear response could be obtained over a signal current range of about 20 db—a brightness ratio of 10 to 1.

The actual reproduced pictures considerably exceeded this range; in fact a brightness ratio of 50 or 100 to 1 was realized. In these tests it was found desirable to hold random interference down about 40 db below the maximum signal, and pattern interference down at least 15 db more.

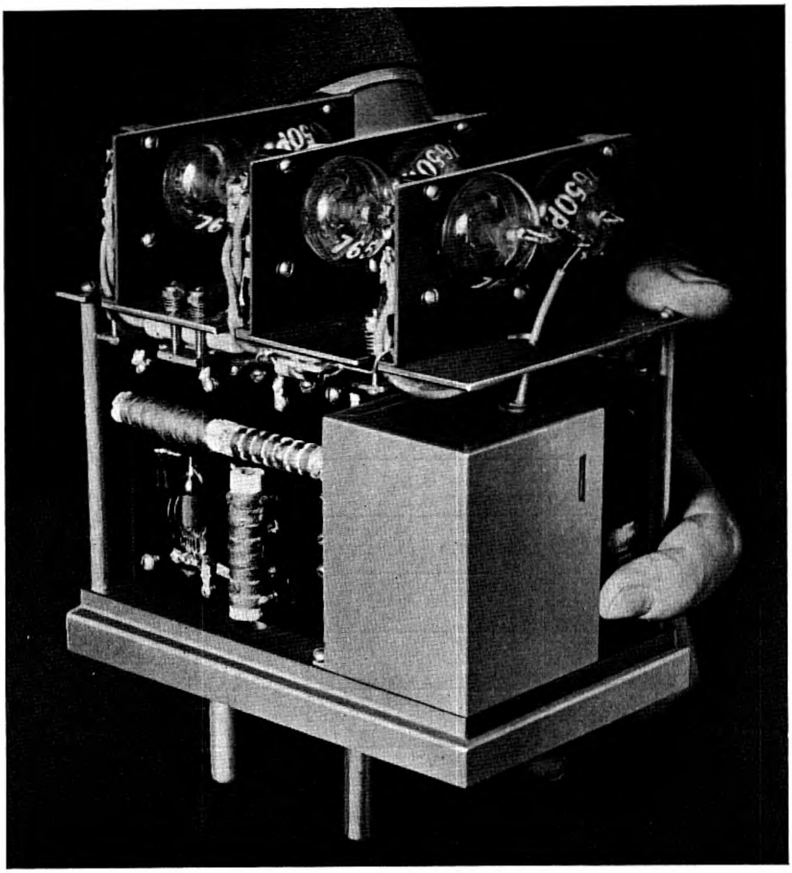


Fig. 16—Photograph of a two million cycle amplifier under development for experiments on the coaxial cable.

The engineers who worked on the system, and outside experts who observed it, expressed the opinion that the reproduced pictures in Philadelphia were substantially the same as those seen on a similar receiving device in New York, thus showing that the cable system itself introduced no appreciable distortion.

Conclusion

As a result of the experimental transmission of the pictures over the coaxial cable from New York to Philadelphia it has been proved that wide-band signals of the type required for television can be satisfactorily transmitted over a coaxial cable system, and that in such transmission the distortion introduced by the wire line circuits can be made so small as to be inappreciable, in its effect on the received picture.

The work on these very wide-band systems has only begun and repeaters and terminal apparatus are now under development capable of handling wider bands of frequency. At the present time work is under way on a two-million cycle system for telephone transmission and a trial installation is being made between New York and Princeton. The system will transmit a frequency band of about two million cycles corresponding to a capacity of 480 telephone circuits. Repeaters on this system will be about 5 miles apart and will consist of unattended boxes somewhat smaller than the one-million cycle repeaters illustrated above and placed either in manholes or on poles along the route. Within these boxes there are placed two amplifiers, one for eastbound transmission, the other for westbound, together with the necessary filters and power supply apparatus. The actual amplifiers themselves are quite small compact units one of which is shown in Fig. 16. Two megacycles, of course, is not a sufficiently wide band to transmit the present R.M.A. standard 441-line television signal, but is a logical step toward more economical telephone circuits. Development is also under way on amplifiers capable of transmitting three megacycle bands of frequency, which should amply satisfy the requirements for transmitting the 441-line television signals now envisioned as standard by the television industry.

References

1. "Systems for Wide Band Transmission Over Coaxial Lines" by L. Espenschied and M. E. Strieby, *Electrical Engineering*, Vol. 53, pp. 1371–1380, October 1934.
2. "A Million-Cycle Telephone System" by M. E. Strieby, *Electrical Engineering*, Vol. 56, pp. 4–7, January 1937.
3. "A Carrier Telephone System for Toll Cables" by C. W. Green and E. I. Green, *Bell Sys. Tech. Jour.*, Vol. XVII, pp. 80–105, January 1938.
4. "Cable Carrier Telephone Terminals" by R. W. Chesnut, L. M. Ilgenfritz and A. Kenner, *Bell Sys. Tech. Jour.*, Vol. XVII, pp. 106–124, January 1938.
5. "Transmission Characteristics of the Coaxial Structure" by J. F. Wentz, *Bell Laboratories Record*, Vol. XVI, pp. 196–200, February 1938.

Stabilized Feedback Oscillators

By G. H. STEVENSON

The author presents a mathematical consideration of the conditions which insure constant frequency of the vacuum tube oscillator under changes of electrode potentials or of the cathode temperature. It has already been shown that the grid and plate resistances may enter into the determination of the frequency. The problem is treated here in the manner suggested in the recent studies of feedback amplifiers. The conditions necessary for stability are developed in terms which are independent of particular circuit configurations and are applicable to certain dissipative circuits as well as to purely reactive systems.

THE frequency deviations that accompany changes of the electrode potentials or of the cathode temperature in many types of vacuum tube oscillators have been recognized for some time as having their origin in the variation of the internal resistances of the tube. Llewellyn has shown [1] that both the grid and plate resistances may enter into the determination of the frequency and, by treating the problem as one of network design, has demonstrated the possibility of making the frequency substantially independent of the tube resistances. He also devised a large number of oscillator circuits stabilized in this way and established the conditions necessary for stabilization in each case.

The problem is treated here in a somewhat more general manner suggested by recent studies of feedback amplifiers.[2] The conditions necessary for stability are developed in terms which are independent of particular circuit configurations and which permit their application to certain types of dissipative circuits as well as to purely reactive systems. While no new fundamental principles are presented, it is thought that the restatement of the known principles in broader terms may be of interest.

The mathematical theory will be developed for the case of the single-tube oscillator circuit, since this is the form most generally used. The extension of the theory to multiple stage circuits presents little or no difficulty. The principal assumptions made are, first, that all of the

---

[1] "Constant Frequency Oscillators," F. B. Llewellyn, *Bell Sys. Tech. Jour.*, January 1932.

[2] "Regeneration Theory," H. Nyquist, *Bell Sys. Tech. Jour.*, January 1932; "Stabilized Feedback Amplifier," H. S. Black, *Bell Sys. Tech. Jour.*, January 1934.

circuit elements except the tube resistances are linear and of lumped character and, second, that modulation effects arising from the nonlinearity of the tube resistances may be neglected. The validity of the second assumption is discussed in the appendix to the article by Llewellyn noted above. Its use permits the treatment of the system as though the resistances were actually linear but variable in magnitude in response to variations of the oscillation amplitude.

THEORY

The essential features of the single-tube feedback oscillator are shown in Fig. 1. The feedback network B is unrestricted in its configuration and complexity and may include the vacuum tube electrode capacitances in addition to the external elements. The impedance system of the tube is reduced to the plate and grid resistances R_1 and R_2 with unilateral coupling between them, the latter being indicated by the inclusion of a generator in series with the plate resistance. The voltage E_1 generated in the plate circuit is proportional to the voltage E_2 between the grid and the cathode and, when the system is oscillating, the latter voltage is produced entirely by E_1 as the result of the coupling through the feedback network.

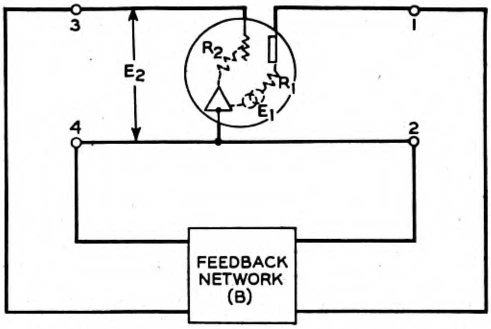


Fig. 1—Elements of a single tube feedback oscillator.

The condition for the existence of self-sustained oscillations is expressed very concisely by the familiar equation

$$\mu\beta = 1, \tag{1}$$

wherein μ and β denote the voltage transfer ratios in the vacuum tube and in the feedback path respectively. The factor μ is the negative of the amplification constant of the tube, the negative sign taking account of the phase reversal inherent in a simple triode. The transfer ratio β represents the ratio of E_2 to E_1 for transmission through the feedback network.

Since the transfer ratios μ and β are both complex quantities, equation (1) expresses the two-fold requirement that the magnitude or modulus of $\mu\beta$ shall be unity and that its phase angle shall be zero. Taking the factors separately, it follows that the modulus of β must be the reciprocal of the modulus of μ and that the phase angles of the two must be equal and of opposite sign. For the single-tube oscillator, the phase angle of β must be 180 degrees since the phase angle of μ has that value.

While the relationships stated above are of simple character, they do not by themselves suffice for the calculation of the oscillation frequency from the constants of the tube and the external circuit. The reason for this is that the values of the tube resistances R_1 and R_2 enter into the determination of the frequency in the general case, and, since these are dependent upon the oscillation amplitude, they cannot be known until the final steady amplitude of the oscillations is known. What happens in an actual oscillator circuit is that, as the oscillation amplitude grows, after initiation, there is a mutual adjustment of frequency and of the resistance values until a condition is reached under which both requirements are met simultaneously. In the case of a stabilized oscillator, since the frequency is independent of the tube resistances, the conditions are simplified and the oscillation frequency can be determined directly by means of the relationships stated above. The non-linearity of the resistances affects only the amplitude of the oscillations.

The evaluation of $\mu\beta$ in terms of the impedance parameters of the circuit permits the determination of the specific circuit conditions in any case for the generation of steady oscillations. The determination is simplified by the consideration that the factor μ has a constant phase angle of 180 degrees so that the variation of the phase of $\mu\beta$ is wholly that of the factor β.

GENERAL FORMULAE FOR $\mu\beta$

The feedback path, or β circuit, is shown separately in Fig. 2, the notations being the same as in Fig. 1. The network B may be of any degree of complexity, but may be assumed to be made up of lumped

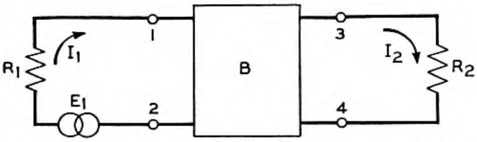


Fig. 2—Simplified schematic of an oscillator feedback circuit.

impedances. In writing down the equations for the mesh currents, let it be assumed that the circuit contains n meshes including the terminal meshes and that the meshes are so chosen that the resistances R_1 and R_2 do not appear as mutual impedances. Designating the meshes in which R_1 and R_2 appear as the first and second respectively, the mesh current equations take the form

$$\begin{array}{ccc|c}
I_1, & I_2, & I_3 \cdots I_n & \\
\cline{1-3}
R_1 + Z_{11}, & Z_{12}, & Z_{13} \cdots Z_{1n} & E \\
Z_{21}, & R_2 + Z_{22}, & Z_{23} \cdots Z_{2n} & 0 \\
Z_{31}, & Z_{32}, & Z_{33} \cdots Z_{3n} & 0 \\
\text{—} & \text{—} & \text{— — —} & \text{—} \\
Z_{n1}, & Z_{n2}, & Z_{n3} \cdots Z_{nn} & 0
\end{array} \tag{2}$$

The subscripts of the Z's denote self and mutual impedances in accordance with the usual conventions, the latter being subject to the reciprocal relationships characteristic of linear systems.

The solution of the above equations for the current I_2 is

$$I_2 = \frac{-\,E\Delta_{21}}{\Delta + R_1R_2\Delta_{11,\,22} + R_1\Delta_{11} + R_2\Delta_{22}}, \tag{3}$$

where Δ is the determinant of the coefficients of equations (2) for zero values of R_1 and R_2, and the other determinants are the minors of Δ obtained by crossing out the columns and rows indicated by the numerical subscripts. Thus Δ_{21} is obtained by crossing out the second column and the first row of Δ and $\Delta_{11,\,22}$ is obtained by crossing out the first two columns and the first two rows.

Since, by definition,

$$\beta = \frac{I_2R_2}{E},$$

equation (3) gives

$$\beta = \frac{-\,R_2\Delta_{21}}{\Delta + R_1R_2\Delta_{11,\,22} + R_1\Delta_{11} + R_2\Delta_{22}}. \tag{4}$$

The factor μ is the negative of the amplification constant of the tube and if the latter be denoted by α, the value of $\mu\beta$ becomes

$$\mu\beta = \frac{\alpha R_2\Delta_{21}}{\Delta + R_1R_2\Delta_{11,\,22} + R_1\Delta_{11} + R_2\Delta_{22}} \tag{5}$$

The determinants appearing in equations (4) and (5) can be expanded by the ordinary processes to give expressions in terms of the mesh impedances in any particular case. However, as they stand, they are significant parameters of the system and, for the present, need no further expansion.

Another general formula for $\mu\beta$ is obtained by making use of the image parameters of the coupling network. If the image impedances at terminals 1, 2, and terminals 3, 4, are denoted by K_1 and K_2, respectively, and the image transfer constant by θ, then

$$\mu\beta = \frac{-\alpha R_2\sqrt{K_1K_2}}{(K_1K_2 + R_1R_2)\sinh\theta + (R_1K_2 + R_2K_1)\cosh\theta}. \tag{6}$$

The two sets of parameters are related by the equations

$$\begin{aligned} K_1^2 &= \frac{\Delta_{22}}{\Delta_{11}}\cdot\frac{\Delta}{\Delta_{11,\,22}}, \\ K_2^2 &= \frac{\Delta_{11}}{\Delta_{22}}\cdot\frac{\Delta}{\Delta_{11,\,22}}, \\ \tanh^2\theta &= \frac{\Delta\Delta_{11,\,22}}{\Delta_{11}\Delta_{22}}. \end{aligned} \tag{7}$$

Equation (6) is useful in many cases because of the fact that the frequency characteristics of the image parameters are well known for a large number of circuit configurations, particularly those of wave filters.

In dealing with many practical oscillator circuits, the simplifying assumption may be made that the coupling network contains only pure reactances. The determinants in equation (5) then become either real quantities or pure imaginaries, thus making it easy to separate the real and the imaginary parts of $\mu\beta$. If the number of meshes in the β circuit, or the number of rows or columns in the determinant Δ, is even, then Δ will be real and if the number is odd Δ will be imaginary. The determinant $\Delta_{11,\,22}$ will be of the same character as Δ, but will take the opposite sign, and determinants Δ_{11}, Δ_{22}, and Δ_{21} will be imaginary when Δ is real and real when Δ is imaginary. Accordingly, equation (5) may be transformed to

$$\mu\beta = \frac{\alpha R_2 D_{21}}{(R_1D_{11} + R_2D_{22}) + j(D - R_1R_2D_{11,\,22})}, \tag{8}$$

in which the D's are determinants of the mesh reactances corresponding respectively to the Δ's of equation (5) having the same subscripts.. The phase angle of $\mu\beta$, denoted by φ, is given by

$$\tan\varphi = \frac{D - R_1R_2D_{11,\,22}}{R_1D_{11} + R_2D_{22}} \tag{9}$$

and the value of $\mu\beta$, when $\tan\varphi$ is zero, by

$$\mu\beta_0 = \frac{\alpha R_2 D_{21}}{R_1 D_{11} + R_2 D_{22}}. \tag{10}$$

The angle φ may be either zero or 180 degrees when $\tan\varphi$ is zero, but which it is may be determined by the sign of $\mu\beta_0$. If this is positive, the phase angle is zero, the phase angle of β being then 180 degrees.

For the simplified case of the pure reactance coupling network the expression for $\mu\beta$ in terms of the image parameters takes different forms depending upon whether the frequency lies in a transmission band or in an attenuating band. At frequencies within a transmission band the image impedances are resistive and the transfer constant θ is a pure imaginary quantity indicating a phase shift without attenuation. Denoting this phase shift by ψ, equation (6) becomes

$$\mu\beta = \frac{-\alpha R_2\sqrt{K_1 K_2}}{(R_1 K_2 + R_2 K_1)\cos\psi + j(K_1 K_2 + R_1 R_2)\sin\psi}, \tag{11}$$

from which

$$\tan\varphi = \frac{(K_1 K_2 + R_1 R_2)}{(R_1 K_2 + R_2 K_1)}\tan\psi. \tag{12}$$

At the cut-off frequencies, equations (11) and (12) become indeterminate. At frequencies in the attenuation bands, the transfer constant θ takes the form

$$\theta = A + j\frac{n\pi}{2}, \tag{13}$$

where A denotes the attenuation and n is an integer the value of which may be different for the different transmission bands of a complex network. The image impedances are pure imaginaries, but their product is real and may be either positive or negative. Simplified forms of equation (6) may be written down for any particular case.

FREQUENCY STABILIZATION

From equation (9), which gives the value of the phase angle of $\mu\beta$, it is at once evident that an essential condition for zero phase angle is that

$$D - R_1 R_2 D_{11,\,22} = 0. \tag{14}$$

Since the quantities D and $D_{11,\,22}$ are functions of frequency, equation (14) determines the frequency or frequencies at which the phase shift is zero and hence determines the oscillation frequency. The equation may be satisfied in two distinctly different ways. In accordance with the first, both D and $D_{11,\,22}$ may be finite and of like sign, in which case the frequency depends upon the resistance product R_1R_2. Since these resistances vary with the oscillation amplitude or with the vacuum tube excitation voltages, a solution of this type is indicative of instability of the frequency.

The second way in which equation (14) may be satisfied depends on the fact that D and $D_{11,\,22}$ may each have zero values at one or more frequencies according to the degree of complexity of the coupling network. If, then, the network can be so designed that D and $D_{11,\,22}$ each have a zero at the same frequency and are of opposite sign elsewhere, the condition expressed by the equation will be satisfied at that frequency for any values of the tube resistances and will be satisfied at that frequency only. Whether or not oscillations can be sustained at the frequency so determined may be ascertained readily with the help of equation (10). Oscillations occurring under the above condition are theoretically stable. As demonstrated experimentally by Llewellyn, a very high degree of constancy of the frequency is obtained in actual circuits.

The method of stabilization described above consists in establishing a limited frequency interval within which the oscillation frequency must necessarily lie and then reducing the width of the interval to substantially zero. The establishment of the finite interval is a matter of the choice of an appropriate circuit configuration and the determination of its limits is effected by suitable proportioning of the elements.

Considered in the light of the image parameters of the feedback network, the method of stabilization consists in making the image phase angle of the network take the value 180 degrees at a frequency within a transmission band. Referring to equations (11) and (12), it will be seen that when ψ, the image phase angle of the network, takes the value 180 degrees, the quantity $\mu\beta$ becomes real and positive, indicating the possibility of self-oscillation. Since the result is independent of the values of the tube resistances, the oscillations are theoretically stable.

In an attenuation band, the transfer constant θ may include a phase angle of 180 degrees which is constant with frequency, but, because of the real component representing attenuation, neither $\cosh\theta$ nor $\sinh\theta$ in equation (6) can become zero at any frequency. To obtain an overall phase shift of 180 degrees in the feedback path it is therefore necessary that

$$K_1K_2 + R_1R_2 = 0, \tag{15}$$

which requires the product K_1K_2 to be negative and, hence, that both image impedances be reactances of the same sign. Since the condition expressed by equation (15) is dependent upon the values of the tube resistances, it follows that frequency stabilization cannot be obtained at frequencies in an attenuation band.

The problem of devising stabilized reactance type oscillator circuits therefore resolves itself into that of obtaining band-pass coupling networks which have phase constants of 180 degrees at frequencies within the pass-bands. Evidently there is a multiplicity of known filter structures that meet the requirements and also many other networks of similar character including all-pass reactance networks. Since each half-section of a filter gives a phase shift of 90 degrees in the pass-band, it follows that the coupling network should be equivalent to at least three half-sections. More complex networks may be used, but with networks equivalent to more than six half-sections oscillations may occur at more than one frequency.

ILLUSTRATIVE EXAMPLES

The principles discussed in the foregoing section will be illustrated by a consideration of the circuit shown in Fig. 3, in which the coupling network is a simple ladder network having four reactive branches. A screen-grid vacuum tube is assumed so that the grid-to-plate capacitance is negligibly small and all feedback is confined to the coupling network. For this circuit, the determinant D has the value

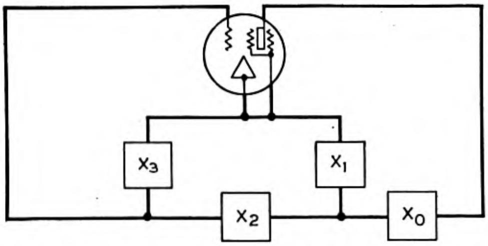


Fig. 3—Oscillator with plate circuit stabilizing impedance.

$$D = X_3(X_0X_1 + X_1X_2 + X_2X_0). \tag{16}$$

The several minors are

$$D_{11,\,22} = X_1 + X_2 + X_3, \tag{17}$$

$$D_{11} = X_3(X_1 + X_2), \tag{18}$$

$$D_{22} = X_0(X_1 + X_2 + X_3) + X_1(X_2 + X_3) \tag{19}$$

and

$$D_{21} = -X_1X_3. \tag{20}$$

The condition that D and $D_{11,22}$ be zero at the same frequency might be met by making the reactance X_3 a simple series resonant combination and proportioning X_1 and X_2 to be resonant at the same frequency as X_3. The two determinants would then both have zeros at this frequency, but oscillations could not occur since D_{21} would also become zero and the feedback would be destroyed. The necessary condition for stabilization is, therefore, that $(X_1 + X_2 + X_3)$ and $(X_0X_1 + X_1X_2 + X_2X_0)$ become zero at the same frequency. It is readily shown that this can be achieved by making the impedance X_0 such that

$$X_0X_3 = X_1X_2 \tag{21}$$

at the frequency for which $(X_1 + X_2 + X_3)$ is zero. The addition of the plate circuit reactance X_0 provides a circuit configuration which makes stabilization possible. The character of the several branch reactances may be such that equation (21) holds at all frequencies but it need hold only at zero of $D_{11,22}$. This minimum restriction permits considerable diversification of the form of the coupling network.

The modified Colpitts oscillator shown in Fig. 4 is a simple case of the general circuit of Fig. 3. For this circuit the reactance determinants have the values

$$D = -\frac{L_0L_2}{\omega C_3}\left(\omega^2 - \frac{1}{L_2C_1} - \frac{1}{L_0C_1}\right), \tag{22}$$

$$D_{11,22} = \frac{L_2}{\omega}\left(\omega^2 - \frac{1}{L_2C_1} - \frac{1}{L_2C_3}\right), \tag{23}$$

$$D_{11} = -\frac{L_2}{\omega^2C_3}\left(\omega^2 - \frac{1}{L_2C_1}\right), \tag{24}$$

$$D_{22} = L_0L_2\left(\omega^2 - \frac{1}{L_2C_1} - \frac{1}{L_2C_3}\right) - \frac{L_2}{\omega^2C_1}\left(\omega^2 - \frac{1}{L_2C_3}\right), \tag{25}$$

$$D_{21} = -\frac{1}{\omega^2C_1C_3}. \tag{26}$$

Both D and $D_{11,22}$ have frequency variations corresponding to those of simple resonant circuits but, in the case of the former, with the sign reversed. The two quantities have the same sign only in the interval between the two resonance frequencies or zeros and, since these resonance frequencies are independently adjustable, the interval may be made as small as may be desired. The interval is reduced to zero and the oscillation frequency stabilized when the elements are so proportioned that

$$L_0 C_1 = L_2 C_3, \tag{27}$$

under which condition the oscillation frequency is determined by the equation

$$\omega_0^2 = \frac{1}{L_2}\left(\frac{1}{C_1} + \frac{1}{C_3}\right). \tag{28}$$

At the oscillation frequency the value of $\mu\beta$ (equation 10) becomes

$$\mu\beta_0 = \frac{\alpha R_2}{R_1 \dfrac{C_1}{C_3} + R_2 \dfrac{C_3}{C_1}} \tag{29}$$

and is equal to unity when

$$\frac{R_1}{R_2} = \frac{C_3}{C_1}\left(\alpha - \frac{C_3}{C_1}\right). \tag{30}$$

Equation (30) can be used to determine the amplitude of the stabilized oscillations if the variation of the resistance ratio with amplitude is known or can be found experimentally. At the moment of inception of the oscillations the amplitude will be infinitesimally small and the tube resistances will generally be such that the initial value of $\mu\beta$ is considerably greater than unity. As the amplitude grows, the plate resistance R_1 tends to increase and the grid resistance to diminish until at a certain amplitude the resistance ratio takes the value given by equation (25). The oscillations then remain steady at this amplitude.

It may be noted that the amplitude relationship (29) holds so long as the capacitances are maintained in the fixed ratio

$$\frac{C_3}{C_1} = \frac{L_0}{L_2} \tag{31}$$

and is independent of their absolute values. If, therefore, the capacitances are varied simultaneously while their ratio is maintained constant, the oscillation frequency will be varied, but frequency stability will be maintained for all adjustments and the oscillation amplitude will remain constant. A similar result may be obtained by varying the inductances simultaneously.

It is instructive to examine the action of the added plate circuit inductance in Fig. 4 in the light of the image parameters of the coupling network. When the values of the inductances and capacitances are unrestricted the coupling network will, in general, have three pass-bands, each pass-band being characterized by a purely imaginary value of the image transfer constant θ corresponding to a phase shift without attenuation. The phase shift is zero at zero frequency, increases by 90 degrees in each pass-band, and remains constant at 90, 180, 270 degrees in the successive attenuation bands. From the expression for tanh θ in equation (7) it follows that the image phase shift of the reactive network will be zero or an even multiple of 90 degrees at the zeros of D and $D_{11,\,22}$ and will be an odd multiple of 90 degrees at the zeros of D_{11} and D_{22}. The general phase shift characteristic of the feedback network in Fig. 4 is shown in Fig. 5. The critical frequencies f_3 and f_4 marking the edges of the second attenuation band correspond to the zeros of D and $D_{11,\,22}$. Frequencies f_2 and f_5 are the zeros of D_{22} and f_1 is the single zero of D_{11}. The pass-bands are indicated by the heavy lines on the frequency scale. An examination of the image impedances of the network will show that in the second attenuation band they are reactances of like sign and are of opposite sign in the first attenuation band.

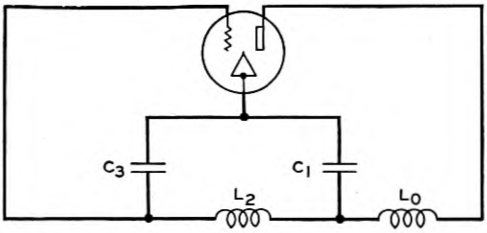


Fig. 4—Colpitts oscillator with plate circuit stabilization.

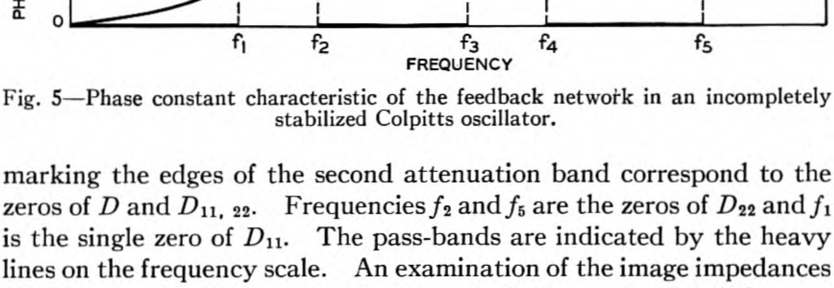


Fig. 5—Phase constant characteristic of the feedback network in an incompletely stabilized Colpitts oscillator.

The necessary condition for self-oscillation stated in equation (15) is satisfied in the attenuation band between f_3 and f_4. The overall phase shift in the feedback path will follow the image phase shift characteristic approximately and will be equal to 180 degrees at some frequency in this range depending upon the values of the terminal resistances provided by the tube space paths. Oscillations may result but their frequency will be unstable. By reducing the width of the attenuation band stability is increased and becomes theoretically complete when the band is reduced to zero. Proportioning the circuit in accordance with equation (27) to give stability makes the two upper pass-bands of the network confluent. If the whole network be proportioned as a one-and-a-half-section constant-*k* filter all three pass-bands become confluent.

Since the stabilization requirement

$$X_0X_3 = X_1X_2$$

need hold only at the oscillation frequency, various possible modifications of the circuit of Fig. 4 become readily apparent. For example, the inductance L_2 may be replaced by a series resonant combination which has the same reactance at the oscillation frequency ω_0. This gives the circuit shown in Fig. 6, in which L_2 is replaced by the combination L_2', C_2' such that

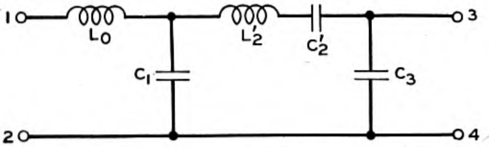


Fig. 6—Modification of the feedback network of the stabilized Colpitts oscillator by the introduction of an extra element.

$$L_2' = L_2 + \frac{1}{\omega_0^2 C_2'}\cdot \tag{32}$$

A further possible modification is shown in Fig. 7 in which L_2 is replaced by a three-element combination comprising a series resonant

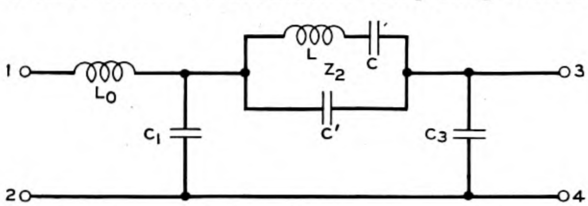


Fig. 7—Further modification of the feedback network of the stabilized Colpitts oscillator.

circuit shunted by the capacitances. At the frequency for which $(X_1 + X_2 + X_3)$ is zero the three-element combination has, as inductive reactance, jX_{20}, which can be computed. The inductance L_0 should then be such that

$$L_0 = \frac{X_{20}}{\omega_0} \cdot \frac{C_3}{C_1}. \tag{33}$$

Evidently the three-element combination in the Z_2 branch is such that it might be replaced by a piezoelectric crystal. To keep the inductance L_0 small, the capacitances C_1 and C_3 should be fairly large so that the resonance of $(X_1 + X_2 + X_3)$ lies close to the crystal resonance.

The foregoing examples are based on the constant-k low-pass filter as a prototype. Evidently high-pass or band-pass filters of the various known kinds might also be used as prototypes and diversified in similar manner. Additional forms may also be found by increasing the number of meshes in the network, but in such cases, the simplest circuits providing frequency stability appear to be the homogeneous single pass-band filter networks. The stable oscillation frequency is the frequency within the pass-band for which the phase constant is equal to 180 degrees. If the network is equivalent to six ladder-type half-sections or more, there will be two or more frequencies for which the phase constant is 180 degrees. Such networks are generally not well-suited for oscillator circuits.

Certain simple configurations which do not admit of complete stabilization in the above manner may be partially stabilized and in actual use may exhibit a very high degree of constancy. The common quartz crystal controlled oscillator with the crystal connected between the grid and cathode of the tube is an example of a partially stabilized circuit. The impedance characteristic of the crystal itself is primarily responsible for the stabilization. Usually the circuit is such as to require the crystal to exhibit an inductive reactance at the oscillation frequency and the impedance characteristic is such that this occurs only in an extremely small frequency interval. The main determinant D for this circuit has a single zero at the resonance of the crystal and complete stabilization would require that the minor $D_{11,\,22}$ have its zero at this frequency also. However, oscillation under this condition would be impossible since the crystal resonance would short-circuit the feedback path and reduce the magnitude of $\mu\beta$ to zero. Actually the zero of $D_{11,\,22}$ lies somewhere in the inductive interval of the crystal at a point fairly close to the resonance frequency. The range in which oscillation is possible is therefore only a fraction of the inductive interval of the crystal and a high degree of stability results.

DISSIPATIVE FEEDBACK NETWORKS

The foregoing sections deal with non-dissipative feedback networks, but the general ideas set forth are applicable to certain types, at least, of dissipative networks. For such networks the transfer constant θ is a complex quantity and may be represented by

$$\theta = A + j\psi, \tag{34}$$

where A denotes the attenuation and ψ the phase constant. When the phase constant is equal to 180 degrees the hyperbolic functions of the transfer constant take the values

$$\begin{aligned} \sinh\theta &= -\sinh A, \\ \cosh\theta &= -\cosh A, \end{aligned} \tag{35}$$

and

$$\tanh\theta = \tanh A.$$

The image impedances will generally be complex, but, if they can be made to become purely resistive at the frequency for which the phase constant is 180 degrees, the value of $\mu\beta$ at that frequency then becomes

$$\mu\beta = \frac{\alpha R_2\sqrt{\rho_1\rho_2}}{(\rho_1\rho_2 + R_1R_2)\sinh A + (R_1\rho_2 + R_2\rho_1)\cosh A}, \tag{36}$$

in which ρ_1 and ρ_2 denote the resistive values of K_1 and K_2. Since this is necessarily a positive real quantity the phase angle of $\mu\beta$ is zero and remains zero independently of variations of R_1 and R_2. Oscillations occurring under this condition are therefore theoretically stable.

TWO TUBE OSCILLATORS

The stabilization of the single-tube oscillator depends on the circumstances that the tube itself produces a constant phase shift of 180 degrees and that feedback networks can be devised to produce a phase shift of this value which is independent of the terminal resistances. Phase shifts of 90 degrees which are independent of the termination can also be provided by means of reactive networks and this property may likewise be made use of in the design of stabilized oscillators. For this purpose it is necessary to have an amplifier which will give a uniform phase shift of 90 degrees over a fairly wide range of frequencies in the neighborhood of the oscillation frequency. A suitable amplifier may consist of two vacuum tubes coupled in tandem by a simple shunt inductance or a simple shunt capacitance. The second tube should be suitably biased to avoid drawing grid current during operation and the coupling reactance should be small in comparison with the plate resistance of the first tube. Preferably the first tube should be of the screen-grid type having a mutual conductance which is independent of the connected output impedance. The two tubes by themselves have a total phase shift of 360 degrees or zero, but the shunt coupling reactance in combination with the internal resistance of the first tube provides a further phase shift of 90 degrees which represents the total effective phase shift of the amplifier.

With a phase shift of 90 degrees in the amplifier, the oscillation frequency will be that for which the feedback path has a phase shift of 90 degrees in the reverse direction. The conditions for frequency stabilization follow readily from the principles already developed.

For a purely reactive feedback network, the expression for $\mu\beta$ may be obtained from equation (8) by substituting $\pm j\alpha$ for the amplification factor. This gives

$$\mu\beta = \frac{\pm j\alpha R_2 D_{21}}{(R_1 D_{11} + R_2 D_{22}) + j(D - R_1 R_2 D_{11,\,22})}, \tag{37}$$

from which is obtained

$$\tan\varphi = \pm\frac{R_1 D_{11} + R_2 D_{22}}{D - R_1 R_2 D_{11,\,22}}, \tag{38}$$

and

$$\mu\beta_0 = \pm\frac{\alpha R_2 D_{21}}{D - R_1 R_2 D_{11,\,22}}. \tag{39}$$

From equation (39) giving the phase angle of $\mu\beta$, it is evident that the phase angle can be zero independently of the magnitudes of the tube resistances only if D_{11} and D_{22} have zero values at a common frequency. This then is the criterion for stability of the oscillation frequency.

In terms of the image parameters of the coupling networks, the value of $\mu\beta$ becomes

$$\mu\beta = \frac{\pm j\alpha R_2\sqrt{K_1 K_2}}{(K_1 K_2 + R_1 R_2)\sinh\theta + (R_1 K_2 + R_2 K_1)\cosh\theta}. \tag{40}$$

If it be assumed that the network is dissipative, the transfer constant has an attenuation component and may be represented by

$$\theta = A + j\psi$$

as in equation (34). When the phase constant ψ has the value ± 90 degrees, equation (40) becomes

$$\mu\beta = \frac{\pm\, \alpha R_2 \sqrt{K_1 K_2}}{(K_1 K_2 + R_1 R_2) \sinh A + (R_1 K_2 + R_2 K_1) \cosh A} . \tag{41}$$

When the feedback network is purely reactive, stabilization of the frequency requires that the phase shift component of the image transfer constant have a value ± 90 degrees within a transmission band. Under this condition the attenuation component of the transfer constant is zero and equation (41) is simplified by the reduction of sinh A to zero and cosh A to unity.

A simple example of an oscillator stabilized in the above manner is shown in Fig. 8. The two vacuum tubes are coupled by a simple shunt inductance of relatively low magnitude and low dissipation. With the high resistance of the screen-grid tube in the first stage this shunt reactance coupling provides a substantially constant phase shift of 90 degrees. A low-impedance transformer might also be used for coupling the stages or, if desired, a four-terminal dissipative network designed to provide the required phase shift in a moderately wide frequency range.

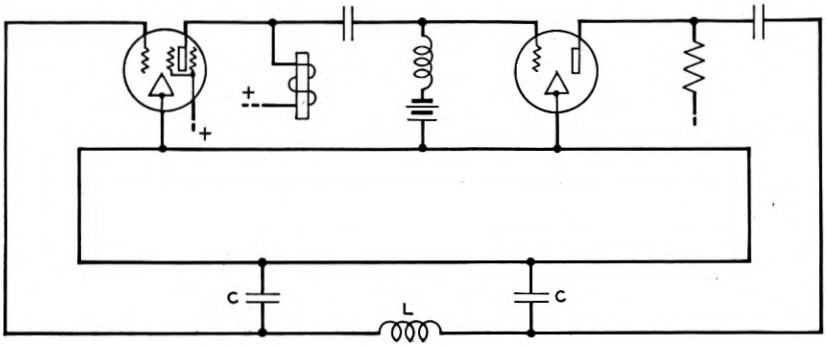


Fig. 8—Stabilized two-stage oscillator.

The feedback network is required to produce a phase shift of only 90 degrees and may therefore have a relatively simple configuration. The direction of the phase shift should, of course, be opposite to that of the amplifier. In the example illustrated the feedback network corresponds to that of a simple Colpitts oscillator. The condition for stabilization is that the two shunt capacitances be equal and the oscillation frequency is that of the resonance of the inductance with one of the two equal condensers.

It will be evident from the foregoing that a very large number of circuit configurations of the feedback network will provide theoretical stabilization of the oscillation frequency either with single stage or with multiple amplifiers. Naturally, not all of these will be of the same practical interest, since an undue complexity of the network may give rise to difficulties in its adjustment and may increase the problem of temperature compensation.

The Discovery of Electron Waves *

By C. J. DAVISSON

THAT streams of electrons possess the properties of beams of waves was discovered early in 1927 in a large industrial laboratory in the midst of a great city, and in a small university laboratory overlooking a cold and desolate sea. The coincidence seems the more striking when one remembers that facilities for making this discovery had been in constant use in laboratories throughout the world for more than a quarter of a century. And yet the coincidence was not, in fact, in any way remarkable. Discoveries in physics are made when the time for making them is ripe, and not before; the stage is set, the time is ripe and the event occurs—more often than not at widely separated places at almost the same moment.

The setting of the stage for the discovery of electron diffraction was begun, one may say, by Galileo. But I do not propose to emulate the gentleman who began a history of his native village with the happenings in the Garden of Eden. I will take, as a convenient starting point, the events which led to the final acceptance by physicists of the idea that light for certain purposes must be regarded as corpuscular. This idea after receiving its quietus at the hands of Thomas Young in 1800 return to plague a complacent world of physics in the year 1899. In this year Max Planck put forward his conception that the energy of light is in some way quantized. A conception which, if accepted, supplied, as he showed, a means of explaining completely the distribution of energy in the spectrum of black body radiation. The quantization was such that transfers of energy between radiation and matter occurred abruptly in amounts proportional to the radiation frequency. The factor of proportionality between these quantities is the ever-recurring Planck constant, h. Thus was reborn the idea that light is in some sense corpuscular.

How readily this circumstantial evidence for a corpuscular aspect of light would have been accepted as conclusive must remain a matter of conjecture, for already the first bits of direct evidence pointing to the same conclusion were being taken down from the scales and meters of

---

* Nobel lecture delivered at Stockholm, Sweden, December 13, 1937 in connection with the award of the 1937 Nobel Prize in Physics by the Swedish Academy of Science to Dr. Davisson and Professor George P. Thompson of London. This lecture was originally published in the year book Les Prix Nobel, 1937.

the laboratory; the truth about light was being wrung from Nature—at times, and in this case, a most reluctant witness.

In an extended examination carried on chiefly by Richardson and K. T. Compton, Hughes, and Millikan, it was brought out that light imparts energy to individual electrons in amounts proportional to its frequency and finally that the factor of proportionality between energy and frequency is just that previously deduced by Planck from the black body spectrum. The idea of pressing the witness on the latter point had come from Einstein who out-plancked Planck in not only accepting quantization, but in conceiving of light quanta as actual small packets or particles of energy transferable to single electrons in toto.

The case for a corpuscular aspect of light, now exceedingly strong, became overwhelmingly so when in 1922 A. H. Compton showed that in certain circumstances light quanta—photons as they were now called—have elastic collisions with electrons in accordance with the simple laws of particle dynamics. What appeared, and what still appears to many of us as a contradiction in terms had been proved true beyond the least possible doubt—light was at once a flight of particles and a propagation of waves; for light persisted, unreasonably, to exhibit the phenomenon of interference.

Troubles, it is said, never come singly, and the trials of the physicist in the early years of this century give grounds for credence in the pessimistic saying. Not only had light, the perfect child of physics, been changed into a gnome with two heads—there was trouble also with electrons. In the open they behaved with admirable decorum, observing without protest all the rules of etiquette set down in Lorentz's manual, but in the privacy of the atom they indulged in strange and unnatural practices; they oscillated in ways which no well-behaved mechanical system would deem proper. What was to be said of particles which were ignorant apparently of even the rudiments of dynamics? Who could apologize for such perversity—rationalize the data of spectroscopy? A genius was called for, and a genius appeared. In 1913 Niels Bohr gave us his strange conception of "stationary" orbits in which electrons rotated endlessly without radiating, of electrons disappearing from one orbit and reappearing, after brief but unexplained absences, in another. It was a weird picture—a picture to delight a Surrealist—but one which fascinated the beholder, for in it were portrayed with remarkable fidelity the most salient of the orderly features which spectroscopic data were then known to possess; there was the Balmer series! and there the Rydberg constant!—correct to the last significant digit! It was a masterpiece. It is important to note that in achieving this *tour de force* Bohr made judicious use of the constant which Planck had extracted from the black body spectrum, the constant h.

It looked at this time—in the year 1913—as if the authentic key to the spectra had at last been found, as if only time and patience would be needed to resolve their riddles completely. But this hope was never fulfilled. The first brilliant triumphs of the theory were followed by yet others, but soon the going became distressingly difficult, and finally, despite the untiring efforts of countless helpers, the attack came virtually to a standstill. The feeling grew that deeply as Bohr had dived he had not, so to speak, touched bottom. What was wanted, it was felt, was a new approach, a new theory of the atom which would embrace necessarily all the virtues of the Bohr theory and go beyond it —a theory which would contain some vaguely sensed unifying principle which, it was felt, the Bohr theory lacked.

Such an underlying principle had been sought for almost from the first. By 1924 one or two ideas of promise had been put forward and were being assiduously developed. Then appeared the brilliant idea which was destined to grow into that marvelous synthesis, the present day quantum mechanics. Louis de Broglie put forward in his doctor's thesis the idea that even as light, so matter has a duality of aspects; that matter like light possesses both the properties of waves and the properties of particles. The various "restrictions" of the Bohr theory were viewed as conditions for the formation of standing electron wave patterns within the atom.

Reasoning by analogy from the situation in optics and aided by the clue that Planck's constant is a necessary ingredient of the Bohr theory, de Broglie assumed that this constant would connect also the particle and wave aspects of electrons, if the latter really existed. De Broglie assumed that, as with light, the correlation of the particle and wave properties of matter would be expressed by the relations:

(Energy of particle) $E = h\nu$ (frequency, waves/unit time).
(Momentum of particle) $p = h\sigma$ (wave number, waves/unit distance).

The latter may be written in the more familiar form $\lambda = h/p$ where λ represents wave-length.

Perhaps no idea in physics has received so rapid or so intensive development as this one. De Broglie himself was in the van of this development but the chief contributions were made by the older and more experienced Schroedinger.

In these early days—eleven or twelve years ago—attention was focussed on electron waves in atoms. The wave mechanics had sprung from the atom, so to speak, and it was natural that the first applications should be to the atom. No thought was given at this time, it appears, to electrons in free flight. It was implicit in the theory that beams of electrons like beams of light would exhibit the properties of waves, that scattered by an appropriate grating they would exhibit diffraction, yet none of the chief theorists mentioned this interesting corollary. The first to draw attention to it was Elsasser, who pointed out in 1925 that a demonstration of diffraction would establish the physical existence of electron waves. The setting of the stage for the discovery of electron diffraction was now complete.

It would be pleasant to tell you that no sooner had Elsasser's suggestion appeared than the experiments were begun in New York which resulted in a demonstration of electron diffraction—pleasanter still to say that the work was begun the day after copies of de Broglie's thesis reached America. The true story contains less of perspicacity and more of chance. The work actually began in 1919 with the accidental discovery that the energy spectrum of secondary electron emission has, as its upper limit, the energy of the primary electrons, even for primaries accelerated through hundreds of volts; that there is, in fact, an elastic scattering of electrons by metals.

Out of this grew an investigation of the distribution-in-angle of these elastically scattered electrons. And then chance again intervened; it was discovered, purely by accident, that the intensity of elastic scattering varies with the orientations of the scattering crystals. Out of this grew, quite naturally, an investigation of elastic scattering by a single crystal of predetermined orientation. The initiation of this phase of the work occurred in 1925, the year following the publication of de Broglie's thesis, the year preceding the first great developments in the wave mechanics. Thus the New York experiment was not at its inception, a test of the wave theory. Only in the summer of 1926, after I had discussed the investigation in England with Richardson, Born, Franck and others, did it take on this character.

The search for diffraction beams was begun in the autumn of 1926, but not until early in the following year were any found—first one and then twenty others in rapid succession. Nineteen of these could be used to check the relationship between wave-length and momentum and in every case the correctness of the de Broglie formula, $\lambda = h/p$ was verified to within the limit of accuracy of the measurements.

I will recall briefly the scheme of the experiment. A beam of electrons of predetermined speed was directed against a (111) face of a crystal of nickel as indicated schematically in Fig. 1. A collector designed to accept only elastically scattered electrons and their near

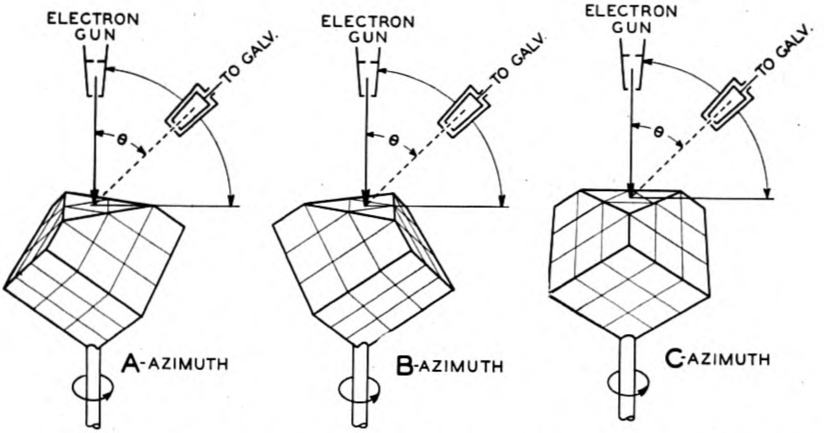


Fig. 1—Schematic diagram showing disposition of primary beam, nickel crystal and collector. Crystal shown revolved to bring one principal azimuth after another into plane of observation.

neighbors, could be moved on an arc about the crystal. The crystal itself could be revolved about the axis of the incident beam. It was possible thus to measure the intensity of elastic scattering in any direction in front of the crystal face with the exception of those directions lying within 10 or 15 degrees of the primary beam.

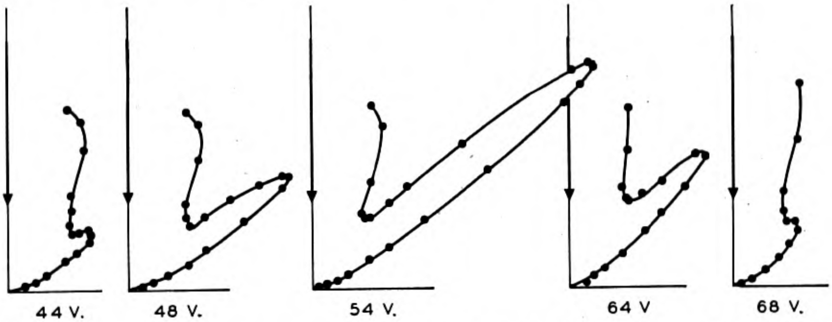


Fig. 2—Polar diagram showing intensity of elastic scattering in A-azimuth (see Fig. 1) as function of latitude angle, for series of primary beam voltages.

The curves reproduced in Fig. 2 show the distribution-in-angle of intensity for a particular azimuth of the crystal. The curves are for a series of electron speeds, therefore, for a series of electron wave-lengths. For a particular wave-length a diffraction beam shines out. Setting the collector on this beam at its brightest and revolving the crystal, the intensity was found to vary in azimuth as illustrated in Fig. 3.

The high peak on the left represents the cross-section-in-azimuth of the beam shown in Fig. 2. Two similar peaks mark the positions of companion beams which with the first form a set of three, as required by the threefold symmetry of the crystal about its (111) directions—the direction of the incident beam. The lesser intermediate peaks are due to a different set of beams which is not here fully developed.

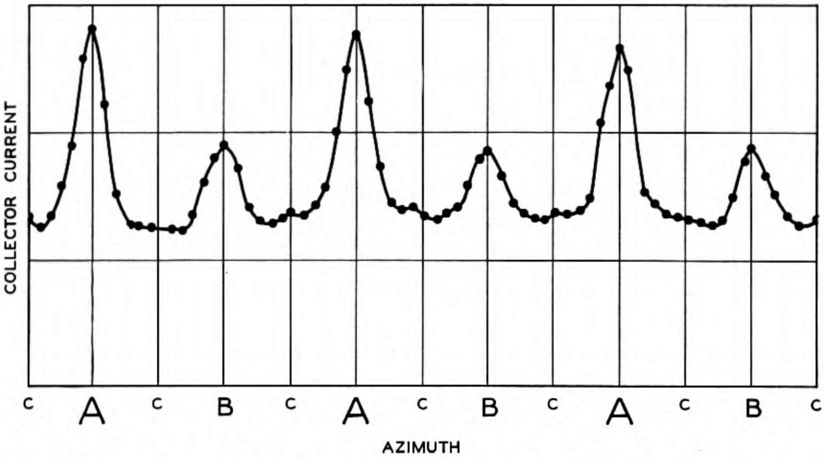


Fig. 3—Curve showing intensity of elastic scattering of 54-volt primary beam as function of azimuth for latitude of peak in 54-volt curve, Fig. 2.

The de Broglie relation was tested by computing wave-lengths from the angles of the diffraction beams and the known constant of the crystal, and comparing these with corresponding wave-lengths computed from the formula $\lambda = h/p$, where p, the momentum of the electrons, was obtained from the potential used to accelerate the beam and the known value of e/m for electrons. If wave-lengths computed from the formula agreed with those obtained from the diffraction data, the de Broglie relation would be verified. How nearly the theoretical values agreed with the experimental is illustrated in Fig. 4. For perfect agreement all points would fall on the line drawn through the origin.

You will realize without my telling you that this series of experiments extending in time over a period of eight or nine years and requiring the construction and manipulation of intricate apparatus was not made by me alone. From first to last a considerable number of my colleagues contributed to the investigation. Chief among these were my two exceptionally able collaborators, Dr. C. H. Kunsman and Dr. L. H. Germer. Dr. Kunsman worked with me throughout the early stages of the investigation, and Dr. Germer, to whose skill and perseverance a great part of the success of the definitive experiments is due, succeeded Dr. Kunsman in 1924.

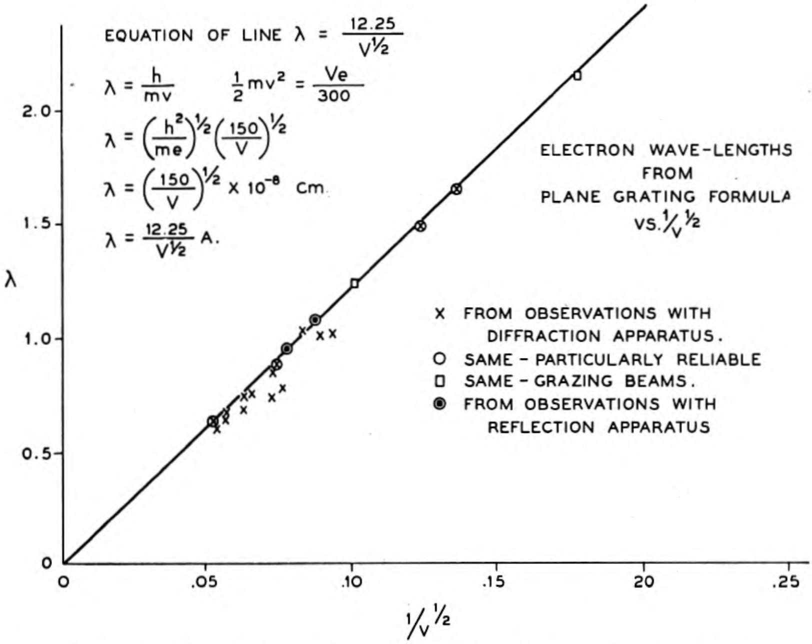


Fig. 4—Test of the de Broglie formula $\lambda = h/p = h/mv$. Wave-length computed from diffraction data plotted against $1/V^{1/2}$, (V, primary beam voltage). For precise verification of the formula all points should fall on the line $\lambda = 12.25/V^{1/2}$ plotted in the diagram.

I would like also at this time to express my admiration of the late Dr. H. D. Arnold, then Director of Research in the Bell Telephone Laboratories, and of Dr. W. Wilson, my immediate superior, who were sufficiently far-sighted to see in these researches a contribution to the science of communication. Their vision was, in fact, accurate for today in ours, as in other industrial laboratories, electron diffraction is applied with great power and efficacy for discerning the structures of materials.

But neither of this nor of the many beautiful and important researches which have been made in electron diffraction in laboratories in all parts of the world since 1927 will I speak today. I will take time only to express my admiration of the beautiful experiments—differing from ours in every respect—by which Thomson in far-away Aberdeen also demonstrated electron diffraction and verified de Broglie's formula at the same time as we in New York. And to mention, as closely related to the subject of this discourse, the difficult and beautifully executed experiments by which Stern and Esterman in 1929 showed that atomic hydrogen also is diffracted in accordance with the de Broglie-Schroedinger theory.

Important and timely as was the discovery of electron diffraction in inspiring confidence in the physical reality of material waves, our confidence in this regard would hardly be less today, one imagines, were diffraction yet to be discovered, so great has been the success of the mechanics built upon the conception of such waves in clarifying the phenomena of atomic and subatomic physics.

Abstracts of Technical Articles from Bell System Sources

Stability of Two-Meter Waves.[1] CHARLES R. BURROWS, A. DECINO and LOYD E. HUNT. The continuous records of the field strength received over a 60-kilometer path on a frequency of 150 megacycles for the year 1936 are analyzed. Preliminary comparison with other paths of the same length indicate that the magnitude of the recorded variations of the signals may be typical of paths of this length.

A reduction in the path length by a factor of two reduced the fading range in decibels by a factor of five.

The results are found to be in agreement with an earlier formula. Fading reduced the field 7 decibels below the average value 1 per cent of the time.

Loudness, Masking and Their Relation to the Hearing Process and the Problem of Noise Measurement.[2] HARVEY FLETCHER. It is shown in this paper how to define loudness and loudness level in a quantitative way. Definite procedures are given for determining experimentally the loudness level of any sound heard by any person. For a typical observer a true loudness scale is developed. The relation of the scale to the loudness level scale is determined experimentally. The scale has been found to be very useful for calculating loudness from the noise spectrogram, the noise audiogram, or the overtone structure of the sound.

The relation between the masking and the loudness produced by a sound has been quantitatively determined and a formula deduced from this relation which has proved useful for calculating the loudness. This formula may be applied with equal success to a normal ear and also to a deafened ear. Evidence has been given that the masking expressed in decibels produced upon any pure tone is equal directly to the agitation of 1.1 per cent of the total nerve endings expressed in decibels above the threshold value for such a patch and at the position where such a tone would be sensed. These loudness relations throw light upon some of the important processes involved in hearing. In particular the data from the masking effects of thermal noise were used to calculate the relation between the position of

---

[1] *Proc. I. R. E.*, May 1938.

[2] *Jour. Acous. Soc. Amer.*, April 1938.

maximum stimulation on the basilar membrane and the frequency of the tone producing the stimulation.

Pick-up for Sound Motion Pictures (Including Stereophonic).[3] J. P. MAXFIELD, A. W. COLLEDGE and R. T. FRIEBUS. Although the basic principles underlying sound pick-up for motion pictures have been understood for some time, the ability to carry them out completely in the presence of the requirements of artistry, photography, lighting, etc., has constituted a difficult problem. The paper discusses some of these problems, particularly with respect to the acoustics of production sets and scoring stages. The problems of stereophonic reproduction are also discussed in some detail.

Practical Application of Telephone Repeaters and Carrier Telephone Systems.[4] J. A. PARROTT. The paper discusses engineering problems in the application of telephone repeaters and carrier systems with which railroad communication engineers recently have been particularly concerned. The first part of the paper deals with crosstalk, noise, balance and overloading considerations in the design of repeatered circuits, particularly from the standpoint of selecting the locations of repeaters to obtain the most satisfactory results on existing lines. The importance of securing test data on the wire facilities to aid in this design work as well as to serve as a guide in improving circuit conditions is emphasized.

The second part of the paper briefly discusses the application of the H1 carrier telephone system and provides transmission data for the preliminary design of the layout of such systems. The Type D and K10 carrier transpositions are described and features of particular interest in their possible use on railroad facilities are discussed.

Sorption of Water by Rubber.[5] R. L. TAYLOR and A. R. KEMP. The effect of several variables on the rate of sorption of water by rubber is discussed. Expressions based on short-time immersion tests are derived which permit calculation of the water content after an extended period of immersion under fixed conditions of temperature and vapor pressure. A sorption coefficient by which one material may be compared with another is suggested, and its application to practical problems is considered.

---

[3] *Jour. S. M. P. E.*, June 1938.
[4] *Proc. Assoc. of Amer. Railroads, Telegraph and Telephone Sec.*, October 1937.
[5] *Indus. & Engg. Chemistry*, April 1938.

Chemical Studies of Wood Preservation—The Wood-Block Method of Toxicity Assay.[6] ROBERT E. WATERMAN, JOHN LEUTRITZ and CALEB M. HILL. Actual decay resistance of treated wood is used as the basis for a simple laboratory technic in the assay of materials advocated for the protection of wood. In its present stage of development the test is a valuable tool in wood preservation studies.

[6] *Indus. & Engg. Chem., Anal. Ed.,* June 15, 1938.

Contributors to this Issue

JULIAN BLANCHARD, A.B., Trinity College (now Duke University), 1905; A.M., Columbia University, 1909; Ph.D., 1917. Professor of Engineering, Trinity College, 1909–1912; Research Assistant in Physics, Columbia University, 1912–1915. Physicist, Research Laboratory, Eastman Kodak Company, 1915–1917; Engineering Department, Western Electric Company, 1917–1925; Bell Telephone Laboratories, 1925–. Dr. Blanchard's work has been concerned primarily with special studies in connection with the development of vacuum tubes and radio.

B. L. CLARKE, B.S., George Washington University, 1921; M.A., Columbia University, 1923; Ph.D., Columbia University, 1924. Bell Telephone Laboratories, 1927–. Dr. Clarke has been in charge of the work in analytical chemistry since 1930.

C. J. DAVISSON, B.Sc., University of Chicago, 1908; Ph.D., Princeton University, 1911; Instructor in Physics, Carnegie Institute of Technology, 1911–17. Engineering Department of the Western Electric Company, 1917–25; Bell Telephone Laboratories, 1925–. As Research Physicist, Dr. Davisson is engaged in work relating largely to thermionics and electronic physics.

In 1928 the National Academy of Sciences awarded the Comstock Prize to Dr. Davisson "for the most important discovery of or investigation in electricity or magnetism or radiant energy" made in this country during the preceding five years, for his work in this field. In 1931 he and Dr. L. H. Germer received the Elliott Cresson Medals from the Franklin Institute, Philadelphia, and in 1935 he received the Hughes Medal of the Royal Society of London.

W. G. GUSTAFSON, B.S. in Electrical Engineering, Union College, 1927; Columbia University, 1929–36. Bell Telephone Laboratories, 1927–. Mr. Gustafson is engaged in work relating to the development of transformers and repeating coils for communication purposes.

W. HERRIOTT was engaged in astronomical research at the Allegheny Observatory from 1914 to 1917. Research in astronomical and aerial photography at the Research Laboratories of the Eastman Kodak Company followed from 1917 to 1920. Between 1920 and 1925 he was engaged in the development of military instruments and of optical apparatus for microscopy, photography and motion pictures at the Bausch and Lomb Optical Company. During the following three years he was in charge of the Scientific Department of the Fairchild Aerial Camera Corporation. In 1928 he joined the engineering department of Electrical Research Products, Inc., coming to the Bell Telephone Laboratories in 1929 to work on optical and photographic problems associated with sound picture apparatus development. In October 1936 he transferred to the Materials Group of the Electromechanical Division of the Telephone Apparatus Development Department.

A. H. INGLIS, B.A., Yale University, 1914. Western Electric Company, Engineering Department, 1914–17. Signal Corps, A.E.F., 1917–19. American Telephone and Telegraph Company, Department of Development and Research, 1919–34; Bell Telephone Laboratories, 1934–. Mr. Inglis has been concerned with both equipment and transmission matters of station apparatus, latterly as Station Instrumentalities Engineer.

W. C. JONES, B.S. in Electrical Engineering, Colorado College, 1913. Western Electric Company, Engineering Department, 1913–25; Bell Telephone Laboratories, 1925–. As Transmission Instruments Director, Mr. Jones is concerned with the development of telephone instruments and similar devices.

H. C. MONTGOMERY, A.B., University of Southern California, 1929; M.A., Columbia University, 1933. Bell Telephone Laboratories, 1929–. Engaged at first in studies of hearing acuity and related problems in physiological acoustics, Mr. Montgomery has been occupied more recently with the study and analysis of speech.

A. E. RUEHLE, B.S., University of Idaho, 1930. Bell Telephone Laboratories, 1930–. Mr. Ruehle's work has been chiefly concerned with applications of the methods of physical chemistry to chemical analysis.

G. H. STEVENSON, B.Sc. in Engineering, University of Glasgow, Scotland, 1906; Instructor in Electrical Engineering, University of Glasgow, 1906–07. Messrs. Barr and Stroud, Glasgow, 1907–11. Western Electric Company, Engineering Department, 1911–24; Patent Department, 1924–25. Bell Telephone Laboratories, Patent Department, 1925–. Mr. Stevenson's work has to do with patent matters in the fields of wave transmission networks and radio transmission systems.

M. E. STRIEBY, A.B., Colorado College, 1914; B.S., Harvard, 1916; B.S. in E.E., Massachusetts Institute of Technology, 1916; New York Telephone Company, Engineering Department, 1916–17; Captain, Signal Corps, U. S. Army, A. E. F., 1917–19. American Telephone and Telegraph Company, Department of Development and Research, 1919–29; Bell Telephone Laboratories, 1929–. Mr. Strieby has been associated with various phases of transmission work, more particularly with the development of long toll circuits. At the present time, in his capacity as High Frequency Transmission Engineer, he directs studies of new and improved methods of carrier frequency transmission over existing or new facilities.